

**CAN ORAL READING FLUENCY SCORES ON DIBELS ORF
PREDICT ISTEP RESULTS?**

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF EDUCATION

BY

GARY L. STORIE

DISSERTAION ADVISOR: DR. WILLIAM SHARP

BALL STATE UNIVERSITY

MUNCIE, INDIANA

DECEMBER 17, 2011

**CAN ORAL READING FLUENCY SCORES ON DIBELS ORF
PREDICT ISTEP RESULTS?**

SIGNATURE PAGE

Dr. William L. Sharp, Department of Educational Leadership, Advisor

Date

Dr. Delbert W. Jarman, Department of Educational Leadership

Date

Dr. Gregory J. Marchant, Department of Educational Psychology

Date

Dr. Paul M. Biner, Department of Psychological Science

Date

ACKNOWLEDGEMENTS

I would like to express sincere gratitude to dissertation advisor and chair, Dr. William Sharp and cognate committee member Dr. Greg Marchant for the time and advice they provided me regarding the completion of this report. The direction and encouragement they provided me was invaluable. I'd also like to thank Dr. David McGuire, Mark Kern, and Shirley Heck, all school administrators who were kind enough to take the trouble and put forth the effort to provide the data for this study. I would like to thank Marty Kuhns, secretary in the Department of Educational Leadership at Ball State University. Without her intervention and assistance with a variety of issues with course work and scheduling I don't think I would ever have completed this degree. Finally, I would like to thank my wife, Linda, an elementary school teacher who has often reminded me about the daily stress of working with children and ensuring they learn. Hopefully, something in this report will provide her and her colleagues, some insight, some motivation, and another tool for instruction that will be helpful in their craft. I also am reminded of my mother, who passed in 1997 and was major influence on developing in me a desire to inquire and learn, and from that give something back to others. This paper is dedicated to her memory.

Gary L. Storie

ABSTRACT

DISSERTATION: Can Reading Fluency Scores on DIBELS ORF Predict ISTEP Performance:

STUDENT: Gary L. Storie

DEGREE: Doctor of Education

COLLEGE: Educational Leadership, Ball State University

DATE: December, 2011

PAGES: 97

The focus of this study was to assess the ability to predict performance on the Indiana Statewide Test for Educational Progress (ISTEP) from oral reading fluency scores on the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Oral Reading Fluency Test (ORF). The scores of both assessments from 306 Indiana students from four different elementary schools were examined. Results of these students' end of year (EOY) ORF scores were compared to their ISTEP scores in third grade in 2008-2009. During that school year ISTEP was administered twice, once in the Fall and once in the Spring. Results from both tests were analyzed using linear, multiple, and logistic regression models. The results showed r values between the ORF scores and the English/Language Arts portion of ISTEP to be .76 (Fall) and .72 (Spring). The correlations between ORF and ISTEP Math were .54 (Fall) and .61 (Spring). Regression coefficients ranged from 1.30 – 1.13 indicating a positive linear relationship between the results of the two assessments. The scores from demographic subgroups were also examined and some differences in the strength of the relationships were found among the groups. The oral reading fluency test was a better predictor for students generally considered at-risk (non-white, disabled, and non-English speaking). Logistic regression

results showed a moderate to strong ability to predict whether or not students would pass ISTEP based on DIBELS scores. After using Beta values to calculate probabilities, it was found a DIBELS ORF score of 100 would provide a strong probability a student would pass ISTEP, both in English/Language Arts and Math. The results imply the need to monitor students' oral reading and provide timely interventions when they struggle, and to focus more on helping children learn to read and less on preparing for ISTEP.

TABLE OF CONTENTS

TITLE PAGE.....	1
APPROVAL PAGE.....	2
ACKNOWLEDGEMENTS.....	3
ABSTRACT.....	4
TABLE OF CONTENTS.....	6
LIST OF FIGURES AND TABLES.....	8
CHAPTER, 1 INTRODUCTION	
Introduction.....	9
Statement of the Problem.....	11
Purpose of the Study.....	12
Significance of the Study.....	13
Research Questions.....	15
Delimitations.....	16
Definitions.....	18
Summary.....	18
CHAPTER 2, REVIEW OF THE LITERATURE	
Overview, DIBELS.....	20
Fluency and Reading Comprehension.....	22
Fluency and the Washington State Test.....	23
DIBELS and Reading Comprehension.....	23
DIBELS and the Florida Test.....	25
DIBELS and the Oregon Test.....	27
DIBELS and the Ohio Test.....	29
DIBELS and the North Carolina Test.....	30
DIBELS and the Pennsylvania Test.....	31
DIBELS and the Iowa Test of Basic Skills.....	32
Instruction and ISTEP.....	32
Reading and Math.....	34
Indiana Results.....	37
Summary.....	38
CHAPTER 3, RESEARCH METHODOLOGY	
Overview.....	40
Research Participants.....	41
Data Collection.....	42
The Instruments.....	44
DIBELS ORF.....	44
ISTEP.....	46
Statistical Model.....	48
Summary.....	50
CHAPTER 4, RESULTS	
Introduction.....	51
Sample Compared to Population.....	54
Descriptive Results.....	55

Normality of the Samples.....	56
Sample Size and Power.....	57
The Relationship Between DIBELS and ISTEP Scores.....	58
Subgroups Analyses.....	62
Model Evaluated.....	66
Pass/Fail.....	68
Summary.....	73
CHAPTER 5, CONCLUSION	
Introduction.....	74
Overall Findings.....	76
The Relationship.....	76
The Subgroups.....	78
Research Questions Answered.....	79
General Summary of the Conclusion.....	84
Discussion.....	85
REFERENCES.....	93
APPENDIXES.....	97

LIST OF FIGURES AND TABLES

Table 1 – PERCENT PASSING ISTEP.....	37
Table 2 – PERCENT OF DEMOGRAHPIC SUBGROUPS.....	42
Table 3 – DIBELS ORAL READING FLUENCY LEVELS OF RISK.....	46
Table 4 – FREQUENCY TABLE.....	53
Table 5 – DESCRIPTIVE STATISTICS.....	53
Table 6 – NUMBER AND PERCENT PASS / DIBELS – ISTEP.....	55
Table 7–SAMPLE NORMALITY RESULTS.....	57
Table 8 – REGRESSION RESULTS.....	58
Table 9 – ANOVA.....	60
Table 10 – COEFFICIENTS.....	61
Table 11 – CORRELATION MATRIX FOR SUBGROUPS.....	64
Table 12 – HIERARCHICAL REGRESSION.....	67
Table 13 – CLASSIFICATION MODEL.....	69
Table 14 – VARIABLES IN THE EQUATION.....	69
Table 15 – MODEL SUMMARY.....	70
Table 16 – CLASSIFICATION TABLES.....	71
Table 17 – VARIABLES IN THE EQUATION.....	72
Table 18 – PROBABILITIES PASSING ISTEP FROM ORF.....	72
Figure 1 – PROBABILITIES OF PASSING ISTEP ENG/LA.....	89
Figure 2 – PROBABILITIES OF PASSING ISTEP MATH.....	89

CHAPTER 1

Introduction

The ability of students to read and comprehend relative to their age and grade in school is a critical concern for educators. Successful academic performance is predicated on whether or not a student reads well. A recent study that examined the relationship between reading ability and subsequent school performance in later grades found students who read at or above level in third grade was a significant predictor of success in 8th and 9th grade course work; that factors leading to better reading also lead to higher graduation rates and better performance in college (Lesnick, et. al., 2010). If this is so, then it makes sense that there may be a connection between reading proficiency and passing the state exam which is prescribed by the No Child Left Behind Act (NCLB) embedded in the Elementary and Secondary Education Act signed into law by President George W. Bush in 2002. In Indiana that test is called the Indiana Statewide Testing for Educational Progress (ISTEP). If there is, indeed, a correlation between reading ability and successful school performance, then there may also be a link between reading proficiency and passing the state exam. As such, then using periodically administered reading assessments to determine reading levels may provide a prediction of students' likely performance on ISTEP. In turn, the data gleaned from reading inventories will impart

important information for teachers to use to help students acquire the skills necessary to achieve a passing score on the state test.

This study targeted third grade students in four different elementary schools over a period of one academic year (2008-2009) and also included second grade reading scores from the spring of the prior school year. In 2008-2009 the State of Indiana Department of Education decided to transition from a fall test to a spring test. This unique circumstance made available two sets of ISTEP test results from the same students to help further validate this study. Student results from reading tests at the end of their third grade year as well as ISTEP results for both the Spring and Fall administrations were collected. The Dynamic Index of Basic Early Literacy Skills (DIBELS) Oral Reading Fluency(ORF) test was used to determine reading fluency levels, which were expressed as levels of risk for achieving or not achieving desired reading proficiency. The primary objective of this study was to determine whether or not end-of-year second grade students' reading levels could predict performance on the third grade ISTEP, including whether or not the students would pass and if there is a correlation between reading level and a student's scale score on the language arts section of the test. Previous research found a connection between reading ability and math performance, therefore DIBELS ORF scores were also compared to Math scale scores to see if there may also be a relationship between the strength of a student's reading fluency and performance on the math portion of ISTEP (Aiken, 1992).

Current policy as a result of recent legislation in the State of Indiana puts the academic proficiency of third grade students in sharp focus. If students at this grade level do not show mastery of reading standards as demonstrated on ISTEP, then these students

may be subject to intensive remedial programs or even retained. Remedial programs can be costly to a school district in a time when economic resources are especially scarce, plus the research on retention indicates an unfavorable outcome in later grades for students who are held back (Jimerson, 2001). Therefore, because of the prevailing attitude among state policy makers and legislators about the crucial aspect of this age regarding reading performance, third grade students were specifically targeted for this study.

A search of the literature revealed studies in six other states where the possibility of using DIBELS ORF to predict students' future performances was researched. Typically strong relationships were found between the ORF and the state tests. The hypothesis for this study claimed a similar relationship between DIBELS ORF scores from the end of the second grade and subsequent results from third grade ISTEP. This study also considered the possibility of a relationship exists between reading fluency and math.

Statement of the Problem

As a result of the No Child Left Behind Act, there is a clear expectation for schools to ensure their students pass the state test, which in Indiana is ISTEP. A lot of pressure is placed upon the principal and faculty to affect student success on the statewide exam, yet the test generally is administered to students only one time per year (2008-2009 year being the only exception). Because of the high stakes nature of the test results, and not knowing until the scores are returned whether students do well or not, contributes to a great deal of uncertainty and anxiety. Schools may face a number of sanctions if students do not do well and fail to meet continuous improvement standards –

a condition for which principals are generally held very accountable. Such sanctions may include the label of a “failing school,” withholding of federal funds, outsourcing remediation services, reorganization of staff, and dismissal of the principal. The problem is worsened because critical information gleaned from test results is provided to the school so infrequently, and the time between test administrations is so long. It is difficult to monitor student learning – the very purpose we have tests in the first place – when the exams are given only one time per year.

Instruction and learning are inherently linked, the progress of which is quantified by assessment. Principals and educators in general need to have a way to tell if their students are attaining skills and knowledge on a timely basis. Of course, it is impractical to administer a large comprehensive summative exam such as ISTEP very often, and once per year should be adequate. However, finding a way to predict how a student is likely to perform on the annual test would be very helpful. Having to wait until the state test scores are returned and analyzed – which in the past was during a time when another school year had begun and the cycle began to repeat itself – renders principals unable to design curriculum, instructional strategies, and remediation for struggling learners. Having a means to predict future performance on the yearly exams would provide valuable information to ensure student success.

Purpose of the Study

The primary goal of this study was to determine if the level of a student’s oral reading fluency ability can predict performance on ISTEP. It was postulated that if a student reads fluently at or above his or her grade level by the end of the second grade, or at the “low risk” designation of DIBELS, then that student will also pass both portions of

the third grade state exam, language arts as well as mathematics. Aligned with the purpose of this study is the hope that if such predictions can be made from students' reading abilities, then educators will have knowledge of what needs to be done to prepare the students adequately for ISTEP. Teachers and administrators trained in elementary education have broad knowledge of the language arts curriculum and the way by which children acquire the ability to read (Salinger et. al., 2010). As such teachers in the early grades can concentrate more fully on reading instruction and not overly worry about test-taking preparations, assimilations, and/or a myriad of other possible assessments the local school or district may require. The main objective of this study was to ascertain the predictability of reading fluency on ISTEP performance and if the conclusion holds true, then there are implications for elementary teachers to have clarity and focus on reading instruction and practice for their students. The short term goal may be to pass ISTEP, but the larger goal is to help children learn to read fluently and comprehend what they are reading.

Significance of the Study

A foremost dilemma of comprehensive summative tests, especially if they are used by the public and governmental agencies to rate a school's overall performance, is the tendency to allow its administration be a single-minded impetus among faculty and administrators. This leads to an over arching emphasis on classroom practices that gear directly toward test-taking readiness, practice testing, and focus on minute content oriented responses. Teachers often complain about the dominant role testing takes in the school and that it detracts from otherwise valid and authentic teaching (Cole, Hulley, Quarles, 2009). Also, students are often subjected to a rigorous and time-consuming

litany of exercises that involve low-level thinking and instructional practices that many believe are simply a matter of “teaching to the test.” In some cases such test prep activities are ethically marginal, ugly, and even illegal (Wright, 2009). Although it is certainly important to have such tests to assess overall learning, it should be the goal of educators to simply teach and ensure the acquisition of learning for its own sake. If this is successful, then positive outcomes on any test should be the end result. It is not the test that is the aim, rather it is student achievement – attaining proficiencies in reading, math, as well as for students to be able think, reason, and use their minds well. “Prepping” for a test really does little of this.

The significance of this study is drawn from the supposition that striving to deploy good teaching methodology should dominate the efforts of educators. Such methods include smaller, more frequent assessments that monitor learning and inform instruction. Therefore, if it can be shown that teachers may have a means to predict student performance on such a high stakes legally required test, in this case ISTEP, by using sound educational practices and formative assessments such as DIBELS ORF, then the classroom can be centered more on true teaching and less on test taking preparations and considerations.

The Elementary and Secondary Education Act (ESEA) was reauthorized in 2002 and is likely to remain for some time, hence annual statewide tests will be with us as well and be given only once per year. The time it generally takes between the test administration and availability of results is usually several weeks and the test results will be used to determine, in part, if a school is making adequate yearly progress (AYP), or making continuous improvement. Due to the critical nature of AYP, it is simply too

important a score to find out only once per year and a month or so after the test is given. Therefore, by finding predictors – or other measures – that will give clues to students’ aptitudes and abilities before, and even between, ISTEP administrations will provide educators powerful tools to monitor learning and take appropriate action and intervention when necessary. Reading is basic to all other academic endeavors so it makes sense that looking at reading ability as a possible predictor of ISTEP performance could prove to be helpful. Teachers can check students’ reading fluency levels often in the course of a school year, and if such levels are slipping, remedial instruction can be provided on a timely basis.

Research Questions

The following research questions were addressed in this study:

- Do third grade students who read at or above the “low risk” designation of DIBELS ORF at the end of their second grade year pass the Language Arts portion of ISTEP in third grade?
- Do third grade students who read at or above “low risk” designation of DIBELS ORF at the end of their second grade year pass the Math portion of ISTEP in third grade?
- Can the number of correct words per minute as scored by DIBELS ORF at the end of the second grade year be used as a predictor of whether or not that student will pass the spring ISTEP (both Language Arts and Math) in the third grade?

- Is the relationship, if any, between the end of the second grade DIBELS ORF and third grade ISTEP scores in the Fall similar to a relationship between these scores in the Spring?
- Is the relationship between DIBELS ORF scores and ISTEP performance similar or different between students from differing demographic subgroups such as gender, race (white, nonwhite), students with disabilities and those without, students in poverty and those who are not, and English Language Learners and those who are not?

Delimitations

Certain variables were under the control of the researcher in this study, the first of which is the size of the sample. The number was planned to be around 300, which included end-of-year second grade 76DIBELS ORF scores and subsequent third grade ISTEP scores. A minimum sample of 105 scores for the regression analysis in this research was needed (Green, 1991). However, to imply greater generalization the number 300 was sought. According to the Indiana Department of Education (2008-2009) website, there were 79,886 third grade students in the public schools. Also, the specific number of elementary schools the researcher included in the study was purposely determined. By obtaining data from four different elementary schools it was hoped an adequate representation of subgroups of students was achieved (i.e., socio-economic status, gender, and race). Third grade students were targeted because of the interest state education policy makers are taking in this elementary level, and the belief they hold that

reading proficiency at the end of third grade is a pre-requisite for later success in schools (Indiana Public Law 109).

The Indiana statewide test, ISTEP, was chosen because it is the annual summative exam that state law requires be administered to all public school students in grades three through eight, which also abides by federal law as stipulated in the Elementary and Secondary Education Act reauthorized in 2002. The Oral Reading Fluency portion of DIBELS was used to determine students' reading fluency abilities because it is a widely used test in schools across the nation (Strauss, 2007) and is an assessment endorsed by the Indiana Department of Education for its Diagnostic Assessment Program, which was implemented in 2011. It has also been determined that fluent readers are more likely to comprehend the printed material they read (Riedel, 2007), and the supposition that strong fluency and subsequent comprehension are essential skills that enable third grade students to pass ISTEP, which is the essential question this study examined.

Although it is informed and guided by norm-referenced data ISTEP is a criterion referenced test. Performance is measured by a scale score and a specific cut score, or a score number the student must achieve to pass. ISTEP scale scores range from 100 to 690. Therefore, the simple question of whether or not a student passes with regard to how well that student did on oral reading fluency was one inquiry for this study. Performances on Language arts, and Math were considered separately since each portion of the test holds its own cut score. Also, the predictability of using DIBELS ORF scores on ISTEP may have been different for Language Arts and Math.

In addition to looking at students' ISTEP performances with respect to cut scores, and whether or not they passed, their actual scale scores were also examined. The

numerical scores (correct words per minute, or cwpm) from DIBELS ORF were compared to scales scores to seek possible predictive correlations. In other words, if a student's DIBELS score was 100 cwpm, then what, if any, was the relationship to that student's score scores on both the Language Arts and Math portions of ISTEP.

Definitions

Definitions that were used for this study include the following:

- ISTEP - Indiana State Test for Educational Progress, to comply with federal law – the No Child Left Behind Act
- DIBELS – Acronym for the Diagnostic Indicators for Basic Literacy Skills, a reading test designed primarily for elementary
- DIBELS ORF (Oral Reading Fluency)– a component of the DIBELS reading test that assesses oral reading fluency
- LOW RISK – Second grade students who achieve a correct words per minute score of 90 are considered “low risk” of struggling with grade level reading
- Scale score – the numerical score on ISTEP that represents as students level of performance on that test
- Cut score – the point at which a student must score at or above to achieve passing on ISTEP

Summary

The purpose of this study was to determine if it is possible to predict students' performance on their third grade ISTEP from end-of-second grade DIBELS ORF scores because it is critical for educators to get information regarding students' academic

progress *before* ISTEP is given. Failing to demonstrate continuous academic progress as measured by ISTEP can result in several sanctions imposed by the government that can adversely affect a school. The correlation between reading fluency proficiency and ISTEP performance was studied because it is theorized that if a student reads at or above the expected level at the end of the second grade, that student will also pass the state exam in the third grade. If it is known a student's reading fluency ability is below the standard, a teacher will be able to intervene with remediation and supplemental instruction throughout the year, and, most importantly, before ISTEP is given. A score from a test such as DIBELS ORF that can be given repeatedly and often, especially in a diagnostic setting, that can also be used to predict ISTEP performance, will be a valuable tool for a teacher to use in planning specific instructional strategies. The essential question this study addressed concerned whether or not a student's reading fluency as measured by DIBELS ORF at the end of the second grade can predict performance on ISTEP in the third grade.

CHAPTER TWO

Review of the Literature

Overview, DIBELS

The fact that learning can often be linked to reading ability is intuitive. To decode letters and sounds and extract meaning from print is a highly complex mental process, one that develops over time and begins at a very young age. Once children are in school an aptitude for reading becomes critical – reading well and achievement are closely related, so much so that if a student struggles with reading the likelihood of that same student graduating from high school is diminished (Lesnick, et. al., 2010). Even poor attitudes held by students about reading can have a negative effect on school performance. Seventy-six fourth grade students were given an Elementary Reading and Attitude Survey (ERAS) and four months later were also given the Indiana Statewide Test for Educational Progress (ISTEP). The researchers found attitude about reading significantly predicts reading achievement on the high stakes test (Martinez, R., et. al., 2008). Certainly teachers in early elementary grades devote a lot of effort teaching reading; much of their day is consumed with literacy lessons and exercises. New legislation in Indiana requires 90 minutes of uninterrupted reading instruction in elementary schools everyday. With an apparent need for students to learn to read well,

and for educators to develop this skill within their students, there is a necessity to frequently monitor learning and provide additional assistance and remediation as needed.

Though there are a number of reading tests and inventories that may be used to assess a student's reading level, the Dynamic Indicators of Basic Literacy Skills, or DIBELS, is one assessment that is widely used across the United States and Indiana for students in the early elementary grades. It attempts to measure the acquisition of emergent reading skills deemed necessary for understanding meaning in print. It is composed of five parts: Onset Recognition Fluency (OnRF – measures phonological awareness), Nonsense Word Fluency (NWF – measures letter-sound correspondence), Letter Naming Fluency (LNF – measures risk for difficulty in literacy), and Oral Reading Fluency (ORF – measures accuracy and fluency with connected text.)

Typically DIBELS is used to ascertain the development of skills of emergent readers (students in grades PK-3). When weaknesses are identified in a student's performance, explicit remedial instruction is provided targeting the specific skills found lacking. For this report, a review of the literature on DIBELS sought to find whether or not results on this reading test, specifically the oral reading fluency component (ORF) which tests reading fluency, can predict future performances on state exams. Eight studies were found that tested such predictive ability of DIBELS. They were "Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading outcomes," by Alysia D. Roehrig, Yaacov Petscher, Stephen M. Nettles, Roxanne F. Hudson, and Joseph K. Torgesen (2007); "The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third grade high-stakes outcomes," by Roland H. Good III, Deborah C. Simmons, and Edward K.

Kame'enui (2001); "The relationship between oral reading fluency and the Ohio Proficiency testing in reading," by Carolyn D. Vander Meer, F. Esward Lentz, and Stephanie Stollar (2005); "Comparing DIBELS oral reading fluency indicator and the North Carolina end of grade reading assessment," by Jeff Barger (2003); "Use of a measure of reading comprehension to enhance prediction on high stakes assessment (2008)," by Edward Shapiro; and "Are fluency measures accurate predictors of reading achievement," by Stephen G. Schilling, Joanne F. Carlisle, Sarah E. Scott, and JiZeng (2007).

Two studies examined the relationship between the DIBELS oral reading fluency test and comprehension. They were "The relation between DIBELS, reading comprehension, and vocabulary in urban first grade students," by Brandt W. Riedel (2007), and "The effectiveness of DIBELS oral reading fluency as a predictor of reading comprehension for high and low income students ," by T. M. Paleologos (2005). Two studies were reviewed that examined the relationship between reading and math. They are, "Evaluation of the relationship between literacy and mathematics skills as assessed by curriculum based measures," by Kristy J. Rutherford (2009), and "Do words count? Connections between mathematics and reading difficulties," by Nancy C. Jordon (2007).

Fluency and Reading Comprehension

According to the National Reading Panel report released in 2000, five variables of development were ascertained for reading proficiency for emergent and young readers. The variables were phonemic awareness (the ability to discern specific sounds), phonics (sound letter relationships), fluency, vocabulary, and comprehension. Of these, a recent body of research has determined that fluency was especially critical for students in the

early elementary grades to attain comprehension from the text they read (Chard, Vaughn & Tyler, 2002). It certainly seems apparent that readers must be able to easily decode the symbols of letters and sounds (phonics) as a prerequisite to processing the meaning of words from the printed page, and to do this effortlessly leaves more cognitive space for seeking understanding (comprehension). If a student can read a passage of grade level equivalent text of 100 words at a 90-95% accuracy rate, that student is considered a fluent reader (Rasinski, 2004).

Fluency and the Washington State Test

The deduction can be made a fluent reader is more apt to comprehend the meaning of words and passages. As such in a high stakes exam setting where printed text is most likely the means by which a student will be posed the questions, the tendency for fluent readers to choose the correct answers largely improves. In a study by Scott A. Stage and Michael D. Jacobson at the University of Washington in 2001 this notion was affirmed. One hundred and seventy-three fourth grade students were given oral reading assessments (not DIBELS), whose scores were then compared to their performance on the Washington Assessment of Student Learning (WASL). The predictive power of the reading assessment score was determined to be .90 – in short, almost all the students who demonstrated fluency on the informal reading assessment met or exceeded the standard set for them on the WSAL.

DIBELS ORF and Reading Comprehension

The oral reading component of the DIBELS assessment measures the ability of a student to identify and read correctly words in a passage of text that is determined to be at a standard for a particular grade level. One of the five domains of reading proficiency as

asserted by the National Reading Panel Report in 2003 is fluency, which this portion of the DIBELS test examines. Just because a student can read words does this also mean that student is able to comprehend the meaning of the text, which is really the goal of reading in the first place? Studies have been done to determine the relationship between fluency and comprehension, and purport that if a student can easily identify and orally call out words successfully, then he also is likely to understand the meaning of the words and aggregated text of which they are a part (Chad, Vaughn & Tyler, 2002, Stage and Jacobson, 2001).

A study by Timon Mark Paleologos at the University of Auburn in 2005 found a positive correlation between fluency and comprehension. His sample included 215 third grade students who took the DIBELS ORF the Stanford Achievement Test (SAT-10) during the 2003-2004 school year. The SAT-10 was used to measure students' reading comprehension abilities. Of the 215 students 112 had a DIBELS score that indicated fluency proficiency, and 103 did not. Overall, Paleolgogos found a correlation between DIBELS ORF scores and subsequent scores on the SAT-10 ($r = .72$), indicating a fairly strong relationship between fluency and reading comprehension (Paleologos, 2005).

In another study reported in *Reading Research Quarterly*, October 2007, Brandt W. Riedel examined the data from 1,518 first graders in an urban setting to determine if DIBELS could be used to predict satisfactory reading comprehension scores. He found that the DIBELS assessment on fluency (ORF) was a strong predictor of reading comprehension. Brandt also examined the other parts of DIBELS and did not find the same predictive ability of how well students would do on subsequent tests.

In his study, Riedel used the Receiver Operating Characteristic (ROC) to determine optimal DIBELS cut scores for predicting satisfactory reading comprehension. According to his findings, the most predictive subset of DIBELS for making such a prediction was the ORF in the middle and end of first grade. Using the model Brandt also found that the various DIBELS subtests differ in their predictive abilities. Additionally he stated, "Adding more subtests to the most predictive subtest does not significantly increase the predictive ability of DIBELS (Riedel, 2007)."

In order to successfully pass the state exam, either the Language Arts portion or Math, students will need to be able, not only to read, but to also understand the text in the test. Knowing precisely what the questions on ISTEP pose and the answer sought is completely a literary exercise, therefore the ability to easily recall the words on the test (fluency) and what they mean (comprehension) is critical. According the results found in the studies by Paleologos and Reidel, the relationship between reading fluency and comprehension is established. The further implication then is if a student can read fluently then his or her performance on the state test should be positively impacted.

DIBELS and the Florida Test

A study was completed at Florida State University in 2007 to examine the concurrent predictive validity of DIBELS ORF with the third grand Florida Comprehensive Assessment Test (FCAT). Researchers sought to answer the question, "Does DIBELS ORF (the oral reading fluency component) predict performance on the state developed reading accountability measure (Roehrig, et. al., 2007)?" The number of participants totaled 32,207 third graders from Florida schools during the 2004-2005 school year. The diversity represented in this sample included 49% female, 36% white,

36% African American, 23% Latino, 3% multiracial, 1.5% Asian, and less than 1% Native American. In addition 75% of the participants were eligible for free or reduced-price lunch, 17% had an Individualized Education Plan (IEP), 12% of the students were limited English proficient, and 3% of the students were considered gifted. The participants were divided into two groups for calibration and cross-validation.

The results of this study indicated a fairly strong correlation between the DIBELS ORF and the FCAT at .71 and .70 over the two administrations (fall and spring). To further help validate this correlation researchers also compared the results between DIBELS and the Stanford Achievement Test (SAT-10), which was administered at the same time and yielded similar findings.

Roehrig and fellow researchers were primarily interested in determining if the use of DIBELS, which is an assessment used multiple times throughout the year for these Florida students, could identify students at risk for failure on the state exam. Such results were ascertained, albeit with what the authors deemed just a reasonable amount of reliability (Roehrig, et. al., 2007).

According to Roehrig, "Current findings indicated that while a student identified as low risk on ORF may reliably be considered truly not at risk, a positive at risk identification on ORF has only moderate consistency in its ability to identify at risk readers. Although the degree of confidence one is able to have in the predictive reliability of ORF is improved using recalibrated scores, the magnitude of the positive predictive power at the Winter 1 and 2 Assessments is below the standard threshold of .75 and limits the accuracy in identifying students who are at risk readers (Roehrig, et. al., 2007)."

The researchers also evaluated differences in results across the demographic subgroups. Roehrig writes, “There was no evidence of predictive bias across several demographic groups in the logistic regression analyses. The most significant predictor risk on the FCAT was ORF, and the interactions between races/ethnicities, levels of socio-economic status, and language status with ORF were not significant contributors to FCAT risk. Hence, the ability of ORF to accurately identify at risk readers of varying demographic characteristics also was not significantly different (Roehrig, et. al., 2007).”

Referencing findings from her study, Roehrig and her colleagues recommended the need to find predictive measures for students’ performances on state tests, in this case the FCAT. DIBELS ORF can be one of those assessments. However the researchers conclude more studies need to be undertaken. The author states, “The FCAT is often regarded among educators as one of the more difficult of the state assessments used to measure grade-level performance in reading. As such, evaluating the utility of DIBELS ORF data for monitoring learning and predicting FCAT outcomes, which serves as the gateway for promotion to fourth grade, is particularly important (Roehrig, et. al., 2007).”

DIBELS and the Oregon Test

In another study entitled, “The importance of decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third grade high stakes outcomes,” by Roland H. Good III, Deborah C. Simmons, and Edward J. Kame’enui in 2001, the predictive ability of DIBELS in relation to the Oregon Statewide Assessment (OSA), was proven with better success than Roehrig. Good III and colleagues found predictive correlations of up to .82 between DIBELS ORF and the Reading Literature component of the OSA. Therefore Good III postulated the utility of

using DIBELS as an assessment to help educators identify literacy weaknesses among elementary students prior to the administration of the OSA in order to provide remedial intervention if necessary.

A portion of this study examined 364 third graders who were given a DIBELS test on reading fluency (OnRF – Onset Recognition Fluency) and later the Oregon Statewide Assessment. Good III found that students who read grade level material at a rate of 110 words per minute or better were likely to meet or exceed expectations on the OSA. In addition, students who read less than 70 words correct per minute on grade level material were not likely to meet expectations on the OSA (Good III, et. al., 2001). There was no mention in this study of subgroup performances, and therefore no disaggregated analyses of the performances of students from different demographic subgroups. However, overall, the results did demonstrate the ability of using DIBELS as an assessment tool that can assist teachers to determine which students to target for additional literacy instruction to further help prepare them for the state exam.

Good III writes, “Our findings consistently indicated that students who scored low on one indicator were at serious risk of not attaining acceptable levels of performance on subsequent measures. For these students, the goal must be ruin the prediction; that is, to alter proactively the instruction and learning conditions sufficiently so that where children began does not forecast where they will end. For this reason, our focus must be on a prevention-oriented assessment and intervention system with utility for making instructional decisions that change student outcomes (Good III, et. al., 2001).”

As with Roehrig, Good III discussed the importance of further research to validate his findings. Good III and his colleagues based this study at the University of Oregon which is also the origin of DIBELS.

DIBELS and the Ohio Test

In 2005 Carolyn D. Vander Meer along with F. Edward Lentz, and Stephanie Stoller did an analysis of third grade students' end-of-year scores on the DIBELS ORF test in comparison with their results on the Reading portion of the Ohio Proficiency Test (OPT). The OPT results were obtained when these students were in the fourth grade. The purpose of her study was to determine if there was a relationship between DIBELS ORF and OPT, which is Ohio's statewide test administered to students pursuant to state and federal legislation. She wrote, "These DIBELS benchmark goals can provide important targets for instruction in third and fourth grade. This study sought to confirm the connection between achieving the DIBELS benchmark fluency goals and passing Ohio's Reading Proficiency Test (Vander Meer, et. al., 2005)."

A total of 364 third grade students, and later fourth grade students participated in the study. Vander Meer and her colleagues used the DIBELS ORF as well as the DIBELS CBM ORF (Curriculum Based Measurement Oral Reading Fluency) components of the reading assessment to determine students' reading levels at the end of their third grade year. The OPT results were obtained when these students were in grade four. The DIBELS ORF goal for reading fluency is 110 words per minute. Of the students who reached or scored above this goal on the ORF as end-of-year third graders, 72% scored proficient or advanced on the reading portion of the Fall OPT as fourth graders. Vander Meer wrote, "It would appear that the third grade benchmark of 110 is

sufficient for establishing a reasonable probability of proficient or advanced levels on the Fall OPT (Vander Meer, et. al.,2005)." In this study, the researchers also found a strong relationship between scores on the DIBELS CBM ORF and the OPT - stating that those students who achieved scores of 93 or higher on the CBM ORF, 89% scored proficient or above on the OPT.

The authors concluded, "In general, with this sample DIBELS performance has an adequate relationship with a standardized test of reading, and benchmark goals and "at-risk" criteria would appear valid for setting goals and deciding which students need interventions (Vander Meer, et. al., 2005)."

DIBELS and the North Carolina Test

Jeff Barger, North Carolina Teacher Academy Trainer, examined the connection between DIBELS ORF and student performance on the North Carolina end of grade test. He stated his conclusion, "The strong relationship between fluency and high stakes testing strongly indicates that ORF does indeed predict performance on high stakes testing (Barber, 2003)." Barger compared the reading scores of 38 third graders to results on the spring exam. The correlation was strong ($r = .73$), which was in line with findings from the similar study by Roehrig, Petscher, Nettles, Hudson, and Torgesen (2007), comparing DIBELS ORF scores to the Florida statewide assessment for third graders ($r = .71$). Barger's results showed that all 26 students who scored 100 correct words per minute on the ORF passed the North Carolina end of grade test.

Barger described the design of his study as more of a correlation than a prediction. "Because," he said, "of the short time between these two measures." Both tests were administered in the spring within a few days of each other. Similar to other

studies seeking to find the predictability of DIBELS scores on subsequent state exams, a regression analysis was used. Though the sample was small ($N = 38$), the finding still seemed consistent with larger studies. Barger wrote in his conclusion, "This study shows DIBELS Oral Reading Fluency measure may be an accurate predictor of whether or not a student will achieve a proficient score on the North Carolina End of Grade Reading Assessment (Barger, 2003)." However, Barger also asserted it was more difficult to predict performances on the NC test of students who scored 69 correct words per minute or lower other than they would be most at risk of not passing the test (Barger, 2003).

DIBELS and the Pennsylvania Test

Edward Shapiro conducted a study at Lehigh University that examined the relationship of scores between DIBELS and the Pennsylvania System of School Assessment (PSSA). The study, entitled "Use of a measure of reading comprehension to enhance prediction on high stakes assessment," was completed in 2008. Participants included 1000 students in grades three to five from six different elementary schools. Like Brandt, Shapiro also used the Receiver Operating Characteristic and regression analyses. The purpose of the study was to see if DIBELS could be used as a screening measure to identify students at risk for not passing the PSSA.

Shapiro concluded a correlation exists – that DIBELS can be used as a valid tool to assist educators to detect students who have learning difficulties in reading comprehension. In his study Shapiro continued with an effort to see if a secondary screening assessment could further enhance the use of DIBELS as a predictor of students' subsequent performances. The 4Sight Benchmark Assessment was used and was also found to have such predictive ability.

DIBELS and the Iowa Test Of Basic Skills

Another study, “Are fluency measures accurate predictors of reading achievement,” conducted by Stephen G. Schilling, Joanne F. Carlisle, Sarah E. Scott, and Hi Zeng was conducted with the same goal – to find whether or not the DIBELS fluency test could predict performance on the Iowa Test of Basic Skills (CTBS). Data were gathered from 44 schools in nine different school districts in the state of Michigan. Students were administered the OnORF in the fall, winter, and spring as well as the ITBS in the spring.

Results of this study showed that the DIBELS test on reading fluency could significantly predict year-end reading achievement scores on the ITBS. It was demonstrated that DIBELS at-risk benchmarks for oral reading fluency were reasonably accurate at identifying second and third grade students who were reading below the 25th percentile at the end of the year on ITBS (80% for second graders and 76% for third graders). In addition, 32% of second grade students and 37% of third grade students who were identified at risk by OnORF benchmarks did not achieve grade level reading comprehension scores on the spring administration of the ITBS.

Reading Instruction and ISTEP

Very little research has been published to date regarding the Indiana Statewide for Educational Progress, or ISTEP. In a search of the ERIC data, base only 15 articles were found, three of which pertain to reading and subsequent performance on the statewide test. One study examined the effect of block scheduling in the middle school on ISTEP performance, another looked at how a basal text reading program might affect ensuing

ISTEP scores, and a third study considered the attitudes of readers and how this correlated to results on ISTEP.

A study by Melanie Beaver in 1998 investigated whether or not the implementation of block scheduling in the eighth grade would help improve the reading results on ISTEP. Her sample data included the ISTEP normal curve equivalent scores (NCE) of reading and language arts from 130 students in 1993 before a block schedule was established. Two years later in 1995, after block scheduling was implemented, 126 NCE scores were obtained from different eighth graders. The premise was additional minutes of reading and language arts instruction as a result of the block scheduling would result in better ISTEP results. By using a t-test to compare the means of the NCE scores and establishing a significance level of .01, it was determined the block schedule affording more time for students to spend practicing and learning literacy skills resulted in ISTEP gains in both language arts and reading (Beaver, 1998).

The effectiveness of a literature-based reading program to improve ISTEP was considered in a study by Anne Burke in 1998. The reading, language arts, and total battery of ISTEP results of third graders from five classes were tracked over a six year period. For the first three years, 1988, 1989, and 1990 the school district used a skills based vocabulary controlled basal text for language arts instruction. In 1991 a new literature based reading program was adopted which continued to be used for the next three years. The ISTEP NCE scores from students in all six years was collected, compared, and subjected to t-test analyses. Though the mean scores fluctuated over the years, no discernible pattern could be detected and no significant effect on ISTEP from the literature based reading program could be determined (Burk, 1998).

In a more recent study by Jeremy Jewell at Southern Illinois University in 2008, scores on an instrument measuring student attitudes about reading and their ISTEP results were compared to find possible predictive correlations. Results on a reading achievement test were also compared to ISTEP. The participants included 151 fourth graders from four different classrooms and were administered the Elementary Reading Attitude Survey (ERAS) during the last two weeks of the school year. The most recent data for the Reading Curriculum-Based Measure (R-CBM), a reading fluency assessment, were taken from archived student records. The ISTEP scores were obtained from this cohort in their fifth grade year. A path analysis was performed on all the data to evaluate interactions among multiple variables. It was found that reading ability and reading attitude significantly predicted reading achievement on ISTEP (Jewell, 2008). “Of note,” wrote Jewell, “is that when reading attitude is increased by 1 standard deviation, ISTEP increased by .22 standard deviations. Correspondingly, when reading attitude increased by 1 point, ISTEP scores increased 2.35 points (Jewell, 2008).”

Reading and Math

In an attempt to bring more consistency and rigor to instruction in classrooms across the state, the Indiana Department of Education recently adopted the Common Core Standards, a nationalized and uniform curricular model. Implementation of these academic standards began in kindergarten during the 2011-2012 school year and will progress to other grades as the years pass. An addendum to the six basic standards entitled “Communication,” requires students to develop and use literacy skills as they learn about mathematics. The addendum states, “The ability to read, write, listen, ask questions, think, and communicate about math will develop and deepen students

understanding of mathematical concepts. Students should read text, data, tables, and graphs with comprehension and understanding. Their writing should be detailed and coherent, and they should use correct mathematical vocabulary. Students should write to explain answers, justify mathematical reasoning, and describe problem-solving strategies. (<http://dc.doe.in.gov/Standards/AcademicStandards/PrintLibrary/docs-math-grade03.pdf>).” The implication infers a strong connection to literacy, suggesting the development of students’ math skills is closely associated to their proficiency in the language arts. As such, it’s plausible there is a relationship between reading and math performances on standardized tests.

In the journal, *The School Psychologist* (2009) Kristi-Becker Rutherford of the University of California stated the research on the relationship between math and reading assessment tools is slim, and what is available centers primarily on students with learning difficulties. Therefore the author undertook a study to determine if oral reading fluency ability and/or reading comprehension could be predictors for math performance. The data included the assessment results from 180 fourth and fifth grade students from an urban Southern California elementary school. The sample included scores from students in general education classroom settings. The research questions were:

1. To what extent does reading, as measured by ORF and Maze, significantly predict performance on applied math performance, as measured by CBM applied math?
2. What measure more strongly predicts applied math performance, Maze or ORF?

A standardized reading curriculum based measurement (R-CBM), for reading fluency, and MAZE, for reading comprehension, were used to test the students' reading abilities. The Monitoring Basic Skills assessment was utilized to test for math. Students were first assigned 25 math computation problems on the MBS, then later given the MAZE, a reading comprehension test, in a group setting. Later, they were individually asked to read aloud a passage from the ORF test while the test administrator recorded the number of words correctly pronounced. Because two independent variables were used, hierarchal regression analyses were performed on the results to determine which reading assessment showed the highest predictive power.

According the results of this study, scores on the MAZE, or reading comprehension test, appeared to be the better predictor of math performance. The author noted this finding was contrary to other studies she reviewed, but that it seems reasonable to think that reading comprehension is more likely to have a stronger effect on math achievement than reading fluency. She also mentioned the smallness of the sample may have contributed to the discrepancy with other research, as well as the fact the other studies primarily focused on students who were learning challenged. In summary, the author alluded to the need for more research on this topic.

In 1972, Lewis R. Aiken, Jr., reported on the effects of literacy on math in the journal, *Review of Educational Research*. He summarized the results of several studies at that time indicating the range of correlations between reading ability and math achievement from .40 to .86. Aiken wrote, "Clearly, understanding the meanings of words in syntax is essential in learning to read all types of materials (Aiken, 1972)." In one particular example Aiken described a study that involved 179 fourth graders split into

two groups. Each group received a pre-test and post-test, with only one group receiving special reading instruction as it pertained to math. On the post test results, the treatment group showed the most gains. From another study, Aiken reported on the correlations between reading vocabulary and arithmetic reasoning on the California Achievement Test (CAT), which showed $r = .71$ for grade three and $r = .69$ for grade four (Aiken, 1972). The author noted the importance of literacy skills to comprehend the concepts of math.

Indiana Results

An overview of the third grade students in Indiana passing ISTEP from 2006 through 2010 reveals a similar passing rate for the Math and English/Language Arts portions of the test (See Table 1).

Table 1

Percent Passing ISTEP in Indiana

Year	<u>Test Component</u>		
	Pct. Passing ELA	Pct. Passing Math	Pct. Passing Both
2006-2007	75	77	71
2007-2008	77	82	73
2008-2009 Fall	72	76	74
2008-2009 Spring	78	65	64
2009-2010	64	68	73

Note. Indiana Department of Education (2011).

Generally the percent of students passing English/Language Arts and those passing Math are within 6 percentage points or so in all given years, however the noteworthy statistic in this graphic is the percent of students who passed both sections. The literature reviewed in this section of the report indicates general correlations between reading and math scores on assessments range from .40 - .81 (Aiken, 1972,) and on the

California Achievement Test, .69 - .71 (Aiken, 1972) suggest the likelihood reading ability has a positive effect on Math achievement. The range of students passing both sections of ISTEP over a period of four years is between 65% and 73%. Though it is not implied these two sets of values are statistically connected, it is inferred students who score well in reading also score well in math. The percent of students passing only one test or the other only may account for academic deficiencies in the part they did not pass.

Summary

If the assumption can be made that fluency helps a student comprehend the meaning of printed text, and if there is a relationship to fluency level and scoring on a standardized achievement measure, then teachers may have a way to predict future performances on such state mandated high stakes tests. There is a variety of fluency assessments, including DIBELS (which is prevalent in schools today), that can be utilized on a frequent basis to determine the reading proficiency of elementary students. Once deficiencies are diagnosed teachers can provide meaningful and timely remedial instruction to better prepare students for the statewide summative assessments, which generally occur only once per year. This practice will give schools that are struggling to meet overall performance goals on the high stakes tests a viable strategy to employ and thereby predict how their students are likely to score – no need to anxiously wait until the once-per-year test results are returned to make informed instructional decisions.

In keeping with the notion that the strength of students' reading abilities has a direct effect on standardized test performance, then the need for effective instructional programs is apparent. Block scheduling that allows more time for students to study reading and language arts resulted in better ISTEP result (Beaver, 1998), and introducing

activities that spur interest and motivation for students to read cause ISTEP scores to rise as well (Jewell, 2008). Hence, the use of DIBELS ORF to monitor students' reading ability and follow-up with necessary adjustments in lesson planning to ensure student progress in reading achievement may lead to better ISTEP scores.

From the studies examined in this literature review, there appears to be a connection between being a fluent reader and achieving a passing score on the state exams in Florida, Oregon, Ohio, North Carolina, Pennsylvania, and Iowa. Therefore it could be hypothesized there may be a similar relationship between DIBELS scores on the ORF and the statewide test for Indiana – ISTEP, or Indiana Statewide Testing for Educational Progress. The rest of this report examines this conjecture.

CHAPTER THREE

Methods

Overview

In order to test the possible utility of using scores on an end-of-year second grade reading test to predict students' subsequent performance on third grade ISTEP, an assessment commonly used in Indiana schools, DIBELS was selected to obtain reading fluency levels. Though DIBELS contains various components, the oral reading fluency test (ORF) was chosen. Studies have shown a strong relationship between fluency and comprehension (Chard, Vaughn & Tyler, 2002; Paleologos, 2005; Riedel, 2007) giving rise to the assumption that if a student reads well then he or she will be able to understand the text and questions on a standardized test such as ISTEP. It is postulated that third grade students in Indiana who are reading proficient at the end of the second grade with regard to fluency will show similar rates of passing their third grade state exam as students did in six other states where this same question was examined. Second grade DIBELS ORF scores were compared to third grade ISTEP results

In Indiana during the school year 2008 and 2009, as part of the transition to a traditionally fall test, there were two administrations of ISTEP – one in September (beginning of the school year) and one in March (closer to the end of the school year). This transition occurred in response to new legislation enacted by the Indiana Legislature

in its previous General Assembly. Therefore, in this unique situation, two sets of ISTEP scores were able to be obtained for the same class within the same school year. DIBELS ORF scores were compared to ISTEP results both in the Fall and Spring. A sample of 306 students' scores from four different elementary schools was collected. The objective was to determine whether or not any relationships existed between end-of-year second grade DIBELS ORF scores and the third grade ISTEP in the Fall, if so identify correlation coefficients and see how they compared to any correlations in the Spring. DIBELS ORF scores were compared to scores on both the Language Arts and Math portions of ISTEP.

This study was broken down into four different examinations: (1) to see if students who scored at “low risk” on DIBELS (which is considered proficient according to test administration guidelines) at the end of the second grade also passed third grade ISTEP Language Arts in the Fall and Spring; (2) to see if students who scored at “low risk” on DIBELS at the end of the second grade also passed third grade ISTEP Math in the Fall and Spring; (3) to see how the end-of-second grade DIBELS numerical scores, stated as the number of “correct words per minute” compared to the third grade ISTEP Language Arts scale scores; and (4) to see how the end-of-second grade DIBELS numerical scores compared to the third grade ISTEP Math scale scores. Regarding any relationships found, further analyses of gender and socioeconomic status were completed as well.

Research Participants

The sample included the end-of-second grade reading fluency scores (DIBELS ORF) and Fall and Spring ISTEP scores of 306 public school third grade students chosen

from a total population in the State of Indiana of 79,886. It was a cluster sample based on grade level and school. The selection of the students was wholly dependent upon the willingness of the principals of the schools to provide the data to the researcher. An attempt was made by the researcher to identify schools that would offer a representation of the demographics for the total population, specifically pertaining to gender and socio-economic status. Two schools represented students from very small rural communities, one school represented students from a small city school, and one school represented a suburban school district.

Table 2 compares the demographic breakdown for students in the third grade in Indiana during the 2008-2009 school year and the sample for the study.

Table 2

Percent of demographic subgroups, state compared to sample

	Indiana	Sample For Study
Male	51	52
Female	49	49
Race (Non-white)	34	35
Poverty (SES)	42	61
Disability	16	14
ESL	5	16

Note. Indiana Department of Education (2011).

Data Collection

The principals from four different elementary schools were contacted about their willingness to provide reading and ISTEP data for this study. The specific information requested included DIBELS ORF scores from end of the second grade administrations, usually given during the month of May or early June. Scale scores from ISTEP from both the Fall and Spring administrations were also obtained and contained both Language

Arts and Math results. The data were organized in a chart listing each student (names omitted) with parallel columns noting the DIBELS ORF scores, the ISTEP Language Arts scale scores for the Fall, the Math scale scores for the Fall, the Language Arts scale scores for the Spring, and the Math scale scores for the Spring. A column was also added to signify a demographic subgroup to which the student belonged. The number “1” indicated male, “2” indicated female, and “3” indicated a student who received assistance for lunch and textbooks.

Pursuant to the Institutional Review Board Guidelines and Family Educational Records Privacy Act (FERPA), the identity of the students was protected and not needed for any part of this study. Principals signed a letter acknowledging their willingness to participate in this study as well as affirming to keep student records confidential. From the data provided by the schools, no means were available to connect any data to specific children. Information was presented in a table format with test scores and other pertinent information without the names of the students disclosed. The data was received hand-delivered or electronically by e-mail.

Once reading test scores and ISTEP scores were received, the data were entered into computer program (SPSS) for regression analyses. Certain scores were disaggregated for the purpose of evaluating possible differences in the results for different demographic subgroups. The student scale scores on both portions of ISTEP were examined with relation to the cut score, which indicated whether or not they passed either or both portions of the test. Results were analyzed and conclusions made about the relationship between the end of the second grade DIBELS ORF and subsequent performance on the third grade ISTEP.

Instruments

The two instruments used in this study included the oral reading fluency portion of DIBELS, or the DIBELS ORF, and the statewide yearly exam administered to all students in grades 3 through 8 in the State of Indiana, ISTEP. DIBELS ORF is a locally administered and scored test and can be given to students any time during the school year. According to its Program Manual it is suggested that students be tested on DIBELS ORF at least once in the beginning of the year and once at the end of the year. With the exception of the school year 2008-2009 ISTEP is administered once per year. Prior to 2008, that administration was always in the Fall (September), after 2009 it became a Spring administration (March), and during the 2008-2009 school year the test was given to all students in Indiana in grades 3 through 8 both in September and March. ISTEP is a standardize test and scored by employees of CTB McGraw Hill, the test's publisher. Results are later provided to the schools.

DIBELS ORF

Several studies have been conducted to determine the reliability and criterion-related validity of DIBELS ORF, which were compiled by the Dynamic Measurement Group (2004). For third grade, DIBELS ORF reliability in a single probe study was found to be .92, and in a multi-probe study reliability was increased to .98. In the same publication studies were cited pertaining to the concurrent validity of DIBELS ORF and various reading tests. A range of correlations were found, from $r = .42$ to $r = .94$ (Dynamic Measurement Group, 2004).

In a more recent study by Catherine T. Goffreda and James Clyde DiPerna at Pennsylvania State University in 2010, a review of the literature on all DIBELS components on validity and reliability was undertaken. Regarding DIBELS ORF Goffreda noted reliability was examined in three peer reviewed articles and one dissertation and that all of the coefficients reported exceeded .80. Seven articles were reviewed with regard to validity and, according to Goffreda, concurrent validity ranged from moderate to high with the median coefficient at .71 (Goffreda, et. al., 2010), which is comparable to the findings in the report on the technical adequacy of DIBELS (The Measurement Group, 2003).

The full array of DIBELS components require teachers to administer the assessments in the classroom setting, with ORF being given to one student at a time. Teachers provide the student a selection of grade level text which students then read orally. It is a timed test and the examiner monitors the quickness of the recall of the words and notes which words are not read or read incorrectly. At the end of the ORF administration, the teacher will score the student's performance as the number of words that were read correctly in one minute, the score actually being expressed as the "correct words per minute (cwpm)." According to DIBELS test administration guidelines, based on the student's score and the range of cwpm that is achieved, the student will be determined to be as "low risk," "moderate risk," or "high risk" for grade level performance in reading. "Low risk" can be considered a fluent reader and "high risk" being the likelihood that a student will struggle with grade level text and reading. For the purpose of this study it is hypothesized that students who are rated as "low risk" on DIBELS ORF will also achieve a passing score on ISTEP Language Arts and Math.

An assumption must be made that teachers who administered the DIBELS ORF for the participants in this study were properly trained and the psychometric integrity of the test results were accurate. However, variation in the subjective nature of a person listening to another person read and assess whether or not particular words are pronounced correctly can lead to some bias, and is possible to have an effect in the final analysis. The quality and sameness of the test administration is one variable in this study that is beyond control of the researcher, and the assumption is made that teachers who assessed their students with DIBELS ORF had reasonable expertise across the board so that the final results a adequate degree of validity.

Table 3

DIBELS Oral Reading Fluency Test – Categories of Risk, 2nd Grade

CWPM Risk	Range
Low	90+
Moderate	70 – 89
High	0 – 79

Note. DIBELS Program Manual (2004).

Indiana Statewide Testing For Educational Progress (ISTEP)

ISTEP is the statewide test for Indiana that complies with federal and state law to assess the educational progress of students in language arts and mathematics. For selected grades, ISTEP also tests student achievement in science and social studies. The test is given once per year to all students in grades 3 through 8, with the school year 2008-2009 being the only exception (ISTEP was given two times – Fall and Spring). Since 2009 ISTEP is administered in the Spring only (March).

The Indiana Department of Education contracts with CTB McGraw-Hill to develop, publish, and score ISTEP tests given each year to Indiana students. The test aligns with the state Department of Education curricular standards for each grade and subject. According to the statistics provided by CTB McGraw-Hill ISTEP has item reliability ranging from .74 to .97 with a mean of .89 for Language Arts, and a range of .79 to 1.00 with a mean of .94 for Math. Test level reliability was reported at .91 for both Language Arts and Math (ISTEP Program Manual, 2009). Regarding validity, CTB McGraw-Hill reported using the Kaiser-Keyer Olkin measure of sampling adequacy for test items and total and sub-group analyses, and indicated results between .96 and .98 for Language Arts and .75 - .86 for Math (ISTEP Program Manual, 2009).

ISTEP for grade three is administered in parts over a three to four day period. For each subtest students are provided 20 to 55 minutes per session, depending on the particular part assigned. ISTEP administration is tightly controlled, with teachers following designated scripts for each subtest when introducing and explaining instructions and time limits to the students. Certain students with IEP's and/or 504 plans are allowed various accommodations including extra time if they need it. According to the ISTEP administration guidelines and ethical considerations, classrooms are to be void of reference materials (on the walls, on desks, or any means available to students), and careful monitoring of test taking activities is required. Once tests are completed the booklets are packaged and sent to CTB McGraw-Hill for scoring. Results are generally provided back to the school within four to six weeks.

There are two major parts of ISTEP – the multiple choice section, where questions are posed and students select an answer from options provided, and the applied

skills section. For applied skills students are to write narratives in their own words in response to prompts and questions in language arts and math. Scoring for the multiple choice section is done electronically; for the applied skills section CTB McGraw-Hill employs and trains scorers to assess tests according to a specified rubric for each item. The uniformity and consistency among scorers regarding the constructed response items on the applied skills portion of ISTEP is largely dependent on the rigor and efficacy of the training provided by CTB McGraw-Hill. The assumption for this study is that the variation between scorers on specific items is minimal, but it is understood this variable is out of the control of the researcher. As with teachers assessing DIBELS ORF, it is a bias, albeit small, that should be considered.

Overall ISTEP performance for each part, Language Arts and Math raw scores are converted into scale scores. Using norm referenced data from pilot tests, the test publisher, in cooperation with representatives from the Indiana Department of Education, establishes a cut score, which is the point on the scale score a student must achieve in order to be considered passing. For the 2008 - 2009 school year the range for Language Arts scores was 100-690, the cut score for passing was 404; the range of scores for Math was 100 - 620 with a cut score at 393. For the purpose of this study, the relationship between second grade end-of-year DIBELS ORF numerical school of “correct words per minute” and students’ third grade ISTEP scale scores was examined.

Statistical Model

The first consideration was to determine if a relationship existed between students’ scores on DIBELS ORF at the end of the second grade year and those same student scores on ISTEP in the third grade. Linear regression was applied to obtain r and

r^2 values, and the associated coefficients, or Beta values, with the DIBELS ORF score being the independent, or predictive variable, and the ISTEP scale score being the outcome, or dependent variable in each case. The raw scores from both DIBELS ORF and ISTEP were entered into SPSS for computation and analysis. To evaluate any possible differences in the ORF-ISTEP relationship among the demographic subgroups, single variable linear regression was run on the scores from each group separately, the results from which were compiled into a large matrix (table) for ease of illustration and comparison. To ensure the application of multiple linear regressions to each group individually was the best model for this study, hierarchical regression was applied as well with the subgroups' scores as covariates.

A research question for this study sought to answer whether or not a student's score on the second grade end-of-year DIBELS ORF could be used to predict that same student's scale score on third grade ISTEP on both the Fall and Language Arts and Math tests. The results from the previous linear regression analyses were used to make this determination. The underlying premise was the number words a student reads fluently would have a relationship with the scale score he or she achieved on ISTEP. Each dependent variable was analyzed separately; linear regressions were run in SPSS multiple times.

Another research question concerned the relationship between the end of the second grade DIBELS ORF and whether or not students passed ISTEP, both in the Fall and Spring of their third grade year. Logistic regression models were used with dummy coding for pass (1) and fail (0). Again, separate SPSS runs were applied to each

dependent variable to determine the constants and Beta values. From these statistics probabilities for passing ISTEP were calculated.

The final research question pertained to a comparison of the outcomes of the regression analyses between Fall and Spring Language Arts and Fall and Spring Math. The results from the linear and logistic regression models were evaluated to make determinations about similarities and differences between the different subject tests and the different seasonal administrations. In order to test the strength and significance of all regression models, appropriate tests of significance were performed on the data.

Summary

The methodology for this study was designed to answer the research questions pertaining to the possible utility of using students' end of second grade DIBELS ORF scores to predict their future performance on third grade ISTEP Language Arts and Math. The sample included test results from 200 students from four different elementary schools. The sample was obtained from principals of these schools, which included data about DIBELS ORF from the end of second grade and ISTEP scale scores from third grade Language Arts and Math from both the Fall and Spring administrations during the 2008-2009 school year. Demographic information was also received regarding the students' gender, race, and economic status. The statistical models included multiple regressions to obtain prediction correlation coefficients between DIBELS ORF and each portion of the ISTEP exam. Also, logistic multiple regressions were run to analyze the relationship between the ORF scores and the Spring and Fall ISTEP results. Such analyses were also applied the subgroups from the different demographic subgroups.

CHAPTER 4

Results

Introduction

It was hypothesized there is a relationship between oral reading fluency as measured by DIBELS ORF, a test given to students at the end of the 2nd grade and these same students' ISTEP scores the following year in the 3rd grade. ISTEP, which is the Indiana statewide test for all students in grades 3-8, includes a Reading/Language Arts portion and a Math portion and was administered twice during the 2008-2009 school year, once in the Fall and once in the Spring. The hypothesis asserts the relationship exists between the ORF scores and results of both Reading/Language Arts and Math parts of ISTEP. As such, this relationship then provides the possibility for predicting future ISTEP performance from students' DIBELS ORF scores. Since predictability was the focus of this inquiry, several regression models were developed to test the overall hypothesis as well as answer the underlying research questions:

- Do third grade students who read at or above the “low risk” designation of DIBELS ORF at the end of their second grade year pass the Language Arts portion of ISTEP in third grade?

- Do third grade students who read at or above “low risk” designation of DIBELS ORF at the end of their second grade year pass the Math portion of ISTEP in third grade?
- Can the number of correct words per minute as scored by DIBELS ORF at the end of the second grade year be used as a predictor of whether or not that student will pass the spring ISTEP (both Language Arts and Math) in the third grade?
- Is the relationship, if any, between the end of the second grade DIBELS ORF and third grade ISTEP scores in the Fall similar to a relationship between these scores in the Spring?
- Is the relationship between DIBELS ORF scores and ISTEP performance similar or different between students from differing demographic subgroups such as gender, race (white, nonwhite), students with disabilities and those without, students in poverty and those who are not, and English Language Learners and those who are not?

To test for an overall relationship between DIBELS and ISTEP scores simple linear regression models were used. Linear, hierarchical, and logistic regression models were used to examine the relationship between DIBELS scores and whether or not students pass ISTEP. From these results observations were also made about the differences and similarities between the subjects, English/Language Arts and Math, and between the seasonal administrations, Fall and Spring (See Table 4).

Table 4

The frequency, percent, and cumulative percent of each subgroup.

Demographic	Frequency	Percent	Cumulative Percent
Gender			
Female	148	48	48
Male	158	52	100
Race			
White	198	64	65
Non-white	108	35	100
Disability			
No	263	86	86
Yes	43	14	100
Econ. Disadvantaged			
No	118	39	39
Yes	188	61	100
Eng. Lang. Learners			
No	257	84	84
Yes	49	16	100

The five variables in this study include DIBELS ORF scores, expressed numerically as the number of "correct words per minute (cwpm)," and scale scores from Fall 2008 ISTEP and Spring 2009 ISTEP. Each ISTEP test contained two parts - English/Language Arts and Math. For DIBELS ORF possible scores begin with 0 words per minute (the lowest being 4) up to the highest score achieved in this sample, which was 205. The scale scores range from 101-960 (E/LA) and 100-620 (Math).

Table 5

Descriptive statistics for each assessment, DIBELS ORF and ISTEP

	N	Minimum	Maximum	Mean	Std. Deviation
ORF	306	4	205	88.77	37.82
Fall Ela	306	201	690	414.21	66.19
Fall Math	306	100	608	394.01	70.44
Spr Ela	306	182	632	427.06	68.16

Spr Math	306	140	696	424.82	81.56
----------	-----	-----	-----	--------	-------

Table 5 shows the minimum score, the maximum score, the mean, and standard deviation for each variable. In this study, DIBELS ORF is the only independent variable.

Sample Compared to Population

For this study, the targeted population included students in Indiana who were in the third grade in the Fall of 2008 and Spring of 2009 and who had taken the DIBELS ORF in the Spring of the previous school year as second graders. In Indiana the total number of third grade students during the school year 2008-2009 was 87,712, making the sample of 306 students slightly more than .3% (three tenths of one percent) of the population. This particular school year was selected for the study because there were two administrations of ISTEP (the only time this occurred in Indiana) with the assumption that statistical analyses concerning both sets of test scores would yield a more informative evaluation of the results. As such, even though the specific targeted population only included third grade students' scores in this particular year, it was hoped implications could be made about whether or not it is feasible to use DIBELS ORF to predict third grade students' ISTEP performance during any year (see Table 2, which shows the comparison of the demographic breakdown of subgroups in Indiana to the sample).

From the sample of scores collected, with regard to how the demographics compare with the State of Indiana, all categories are comparable except "poverty," or those students who were eligible for financial assistance for meals and textbooks. To control for this discrepancy, linear regression analyses were run on all subgroups that

would yield comparisons in order to discern whether or not belonging to a particular demographic has an impact on the predictability of DIBELS ORF on ISTEP.

Descriptive Results

The following table depicts the basic results of the DIBELS ORF and ISTEP scores including each administration of the test and both portions (Reading/Language Arts and Math). On DIBELS ORF to be considered "low risk" for reading fluency, the student must achieve the standard of correctly reading at least 90 words per minutes. For ISTEP, a cut score of 404 or more must be achieved in order to pass the English/Language Arts test, and a cut score of 393 or more must be achieved to pass the Math test.

Table 6

Number and percent of students at who scored at or less than 90 cwpm on DIBELS ORF for "low risk" designation and ISTEP Pass/Fail

Assessment	Number	Percent
DIBELS ORF		
Low Risk	163	53
Less than Low Risk	143	47
ISTEP PASS		
Fall 2008		
Eng/LangArts	188	61
Math	176	58
Spring 2009		
Eng/LangArts	218	71
Math	222	73

From this information it can be seen that more students passed both portions of ISTEP during both test administrations than did achieve a score of 90 correct words per minute on DIBELS ORF ("low risk" designation, or reading at level). However, questions with regard to correlations between ORF and ISTEP as well as using ORF to

predict ISTEP performance remain and are addressed later in this study. Obviously there are many factors that account for student achievement, with reading fluency being considered as one of them. Of the 163 students who scored 90 or more correct words per minute on ORF, 99 of those students also passed both portions (four sections) of ISTEP on both the Fall and Spring administrations. An additional 16 students who scored 90 or more on ORF passed three of the four sections of ISTEP. Overall, with regard to ISTEP only, students performed much better during the Spring tests than they did in the previous Fall.

Normality of the Samples

To better infer generalizations to the population a sample where the data are normally distributed improves the strength of any deductions made. The conjecture the five variables used in this study met the assumption of normality was examined. The following table summarizes the normality calculations of each variable, the skew and kurtosis values and the associated standard errors.

Table 7

Sample Normality Test Results

Variable	ORF	Fall Ela	Fall MA	SprEla	Spr MA
Skewness	-.04	.04	-1.04	-.70	-.53
Error	.14	.14	.14	.14	.14
Kurtosis	-.16	1.48	2.55	1.82	1.51
Error	.28	.28	.28	.28	.28

From these calculations the sample data appear to deviate from normality, especially ISTEP scores from Fall Reading/Language Arts and Math and ISTEP scores from Spring Math. Deviations for the other scores, including DIBELS ORF show values

less than 1, which suggest a closer fit to normality albeit with some degree of skewness and kurtosis. Outliers in the data may account for portions of the abnormalities, including two scores of 4 and one score of 205 on ORF, two scores of close to 700 on ISTEP Fall E/LA, and three scores of over 600 on ISTEP Spring Math.

The goal of being able to make general references to the population as a result of this study is somewhat compromised by the lack of normality of the data, however some generalizations might be appropriate. Jeremy Miles and Mark Shevlin write in their book, *Applying Regression and Correlation: A Guide for Students and Researchers* (2004), “The problem is that we are not really interested in whether the distribution is significantly different from a normal distribution (using significant in the technical sense of the word). We are interested in knowing whether the distribution is sufficiently skewed that it matters.” Miles and Shevlin provide a “rule of thumb,” and write, “If the skewness statistic is less than 1.0 there should be little problem. If the skewness is greater than 1.0 but less than 2.0, you should be aware that it might be having an effect on your parameter estimates, but that is probably OK. Finally if the skewness statistic is greater than 2.0 you should begin to be concerned (Miles & Shevlin, 2004).”

The lack of normality in the sample needs to be considered with the findings of this study along with the notions put forth by Miles. None of the normality statistics exceeded 2.0, however some were between 1.0 and 2.0 providing some reason for concern. The outliers may be part of the explanation, but such exceedingly high and low scores are likely to exist in the population as well.

Sample Size and Power

In order to test the overall fit of the regression models used in this study as well as be able to test the predictor variable (DIBELS ORF), a minimum sample size of 105 was needed (Green, 1991). The sample for this study was comprised of 306 student scores from DIBELS ORF at the end of their second grade year, and these same students' ensuing ISTEP scores as third graders, therefore exceeding the minimum required according to Green. With regard to effect size, a sample with this amount of cases should yield a large effect (Field, 2009; Miles & Shevlin, 2004). Also, according to Miles, the sample used for this study and its estimated effect size should provide a strong degree of power (80% chance) of finding a relationship between DIBELS ORF and ISTEP should that relationship exist in the population (Miles & Shevlin, 2004).

The Relationship Between DIBELS and ISTEP Scores

To evaluate the overall predictability of DIBELS ORF 2nd grade end-of-year (EOY) scores on ISTEP performance the following year (3rd grade) for both the Fall and Spring administrations, linear regression analyses were performed multiple times. The independent variable in each case was the ORF score; the four dependent variables, tested one at a time, included students' scale scores from ISTEP for both the English/Language Arts (Ela) and Math portions of the tests. The following table summarizes the findings.

Table 8

Regression results for DIBELS ORF predicting ISTEP scores

	<i>r</i>	<i>r</i> Square	Adjusted <i>r</i> Square	Std. Error of the Estimate
Spring 2009				
Ela	.72	.52	.52	47.17
Math	.54	.29	.29	68.94
Fall 2008				
Ela	.76	.57	.57	43.50

Math	.61	.37	.37	56.01
------	-----	-----	-----	-------

From these results it is determined there is a relationship between student's ORF scores and their subsequent performance on ISTEP for both English/Language Arts and Math. The r reported, .76 for E/LA Fall 08 and .72 for Spring 09, values which are analogous to Pearson r (Field, 2009), indicate a fairly strong correlation. r^2 shows that over 50% of the variance in this model for E/LA ISTEP results can be explained by students' oral reading fluency as measured by DIBELS ORF (Fall 2008 E/LA $r^2 = .57$, and Spring 2009 R/LA $r^2 = .52$). Based on the r^2 values of the ISTEP scores, this model appears to be a good fit for the data (Field, 2009).

Regarding students' ISTEP Math performance and their previous scores on DIBELS ORF, a relationship is also present, however not as quite strong as English/Language Arts. Correlations, as indicated by r show values of .61 for Fall 08 Math and .54 for Spring 09 Math. The r^2 findings denote about 37% of the difference among Math scores in the Fall of 08 and about 28% of the differences among Math Spring 2009 scores can be explained by oral reading fluency scores the previous year. From these results, unlike the r and r^2 values for both Reading/Language Arts administrations Fall and Spring, which are very close, the values for ISTEP Math are considerably different. DIBELS ORF scores appear to be a better predictor of Fall ISTEP Math than Spring Math, however Oral Reading Fluency, as measured by DIBELS as the number of "correct words per minutes," can be considered a stronger predictor of students' performance in Reading/Language Arts than Math. Essentially, about half of

the in the variance in ISTEP E/LA can be explained by oral reading fluency, for Math the effect is less.

With regard to regression results, an analysis of variance (ANOVA) for each dependent variable (ISTEP scores) was examined for the probability of observing a value greater than or equal to F that is less than 0.01. In this model the significance statistic for Table 9

Analysis of variance result showing F statistic, $sig < .01$ for each dependent variable

IV	Model	Sum of Squares	df	Mean Square	F	Sig.
Spring 2009						
Ela	Regression	740803.38	1	740803.38	332.99	.000
	Residual	676313.56	304	2224.72		
	Total	1417116.94	305			
Ma	Regression	583932.63	1	583932.63	122.85	.000
	Residual	1444976.49	304	4753.21		
	Total	2028909.11	305			
Fall 2008						
Ela	Regression	761044.46	1	761044.46	402.27	.000
	Residual	575128.16	304	1891.87		
	Total	1336172.61	305			
Ma	Regression	559660.93	1	559660.93	178.41	.000
	Residual	953623.02	304	3136.92		
	Total	151283.95	305			

the F -test indicates there is essentially no chance (less than one in 1,000) that the observed correlations between the ORF scores and ISTEP shown previously is solely due to random sampling error (Table 9). For each set of ISTEP results, Fall 2008 and Spring 2009 $sig = .000$ and $p < .001$. Therefore the null hypothesis, that there is no relationship between DIBELS ORF and each set of ISTEP scores, can be rejected.

The slopes of the regression lines were calculated in each case with a separate analysis of each dependent variable. The ORF scores established the X (the independent variable) axis and ISTEP scores the Y (the dependent or outcome variable) axis.

The regression coefficients were determined for each outcome variable (See Table 10).

Results displayed include the constant, where the regression line intercepts the Y axis or the likely score on ISTEP without the consideration of ORF reading scores, and the slope, or regression coefficient, expressed as the Beta values. Other indicators, the t value

Table 10

Regression results for constants and coefficients

Assessments	Unstandardized Coefficients		Standardized Coefficients	<i>t</i>	Sig.
	<i>B</i>	Std. Error	Beta		
Spring 2009 Ela					
Constant	311.39	6.89	.72	45.20	.000
ORF	1.30	.07		18.25	.000
Spring 2009 Math					
Constant	322.12	10.07		31.99	.000
ORF	1.16	.10	.54	11.08	.000
Fall 2008 Ela					
Constant	296.97	6.35		46.75	.000
ORF	1.32	.07	.76	21.06	.000
Fall08 Math					
Constant	293.47	8.18		35.88	.000
ORF	1.13	.09	.61	13.36	.000

and sig show the results of these analyses to be statistically significant at $p < .01$. It is determined by these regression analyses that there is a predicted rise of 1.3 points in both 3rd grade Fall and Spring Language Arts scale scores for every 1.0 points, or one additional "correct word per minute," score on DIBELS ORF at the end of the 2nd grade. The prediction is similar for Math, showing Beta values of 1.2 (Spring) and 1.1 (Fall).

Accordingly, referring to an example from the data set used in this study, a particular student scored 63 correct words per minute on DIBELS ORF also had a scale score of 382 on Fall English/Language Arts, which failed to meet the cut score for passing at 404. From this model of regression analysis, it is inferred had that student read fluently on DIBELS an additional 17 words (to achieve the “low risk” designation of the ORF at 90), the prediction is that the E/LA scale score would have risen an additional 17 points to reach 404 (1.303×27 , rounded), which would have been sufficient for passing. Certainly this deduction is made from the regression coefficient applied to a particular raw score and there is the consideration of the error between predicted scores and actual scores, however this model does suggest a fairly significant relationship between DIBELS ORF and ISTEP; the slope of the regression line is not flat, it does rise (as ORF scores increase so do ISTEP scale scores).

Subgroups Analyses

The sample data was segregated into subgroups for regression analyses according to the following demographics: Gender (male/female), race (white/nonwhite), socioeconomic status (yes/no), disability (yes/no), English Language Learners (yes/no). Simple linear regression, or correlation, was calculated for each subgroup independently for simple comparisons of the r , r square, intercept (constant), and slope (regression coefficient) of each demographic. DIBELS ORF scores always served as the independent variable after being sorted by classification, and the dependent variables were the ISTEP results of each group on each administration and portion of the test. The main objective was to determine differences for each category of values for each

subgroup without consideration of possible interactions between demographic classifications (see Table 11).

With a focus only on the statistics regarding ORF and ISTEP English/Language Arts, when linear regression is applied to each of the subgroups scores a number of differences in the results emerge. Generally the scores pertaining to gender (male and female), students without disabilities, and English speaking students show a comparable relationship between DIBELS ORF and ISTEP to the overall group results. Though slightly higher, surprisingly the regression calculations of students eligible for lunch and textbook assistance (SES), a marker for poverty, are similar as well for English/Language

Table 11

Correlation Matrix for Subgroups

Group	Test	<i>r</i>	<i>r square</i>	Intercept	Beta
Male (N = 158)	Fall ELA	.77	.59	294.54	1.34
	Fall MA	.62	.38	304.62	1.09
	Spring ELA	.75	.56	301.26	1.41
	Spring MA	.55	.30	335.22	1.14
Female (N = 148)	Fall ELA	.73	.54	300.30	1.30
	Fall MA	.61	.38	278.11	1.21
	Spring ELA	.69	.48	324.16	1.17
	Spring MA	.55	.30	303.66	1.23
White (n = 198)	Fall ELA	.70	.49	317.56	1.20
	Fall MA	.53	.27	325.32	0.90
	Spring ELA	.66	.43	339.65	1.09
	Spring MA	.42	.18	369.15	0.80
Non-white (n = 108)	Fall ELA	.80	.64	284.14	1.26
	Fall MA	.65	.42	264.84	1.28
	Spring ELA	.77	.59	285.02	1.46
	Spring MA	.59	.35	278.29	1.42
SES – No (n = 118)	Fall ELA	.67	.45	333.13	1.10
	Fall MA	.49	.24	342.81	0.73
	Spring ELA	.60	.36	364.10	0.91
	Spring MA	.41	.17	376.03	0.79
SES – Yes (n = 188)	Fall ELA	.76	.58	288.95	1.32
	Fall MA	.62	.39	275.07	1.30

Disability – No (<i>n</i> = 263)	Spring ELA	.74	.55	291.62	1.42
	Spring MA	.53	.28	307.66	1.21
	Fall ELA	.73	.54	299.41	1.30
	Fall MA	.60	.36	296.13	1.11
Disability – Yes (<i>n</i> = 43)	Spring ELA	.69	.48	317.89	1.24
	Spring MA	.53	.28	321.02	1.17
	Fall ELA	.83	.68	290.11	1.35
	Fall MA	.60	.36	283.93	1.24
ELL – No (<i>n</i> = 257)	Spring ELA	.81	.66	287.62	1.57
	Spring MA	.51	.26	327.14	1.08
	Fall ELA	.74	.55	304.42	1.29
	Fall MA	.58	.34	307.61	1.03
ELL – Yes (<i>n</i> = 49)	Spring ELA	.70	.49	321.24	1.23
	Spring MA	.50	.25	344.47	1.06
	Fall ELA	.85	.73	274.69	1.29
	Fall MA	.70	.49	242.27	1.47
	Spring ELA	.82	.67	275.92	1.53
	Spring MA	.64	.42	276.95	1.43

Note. For all correlations *F* significant at $p < .01$; all coefficients *t* significant at $p < .01$.

Arts ISTEP scores ($r = .76$ Fall, $r = .74$ Spring). For the total group, including all students, these English/Language Arts values are $r = .72$, (Fall) and $r = .76$ (Spring). In short, between 50-59% of the ELA ISTEP score can be explained by performance on ORF for students regardless of gender and poverty status. For this study students included in the sample who had an Individualized Education Plan (IEP) were designated as learning disabled regardless of the special education label they held. Looking only at regression statistics between ORF and ISTEP English/Language Arts (Fall and Spring) there are significant differences in these relationships between students with disabilities and those without. For students not labeled disabled $r = .73$ (Fall) and $r = .69$ (Spring), which is similar to the results of the whole group (the Spring correlation is somewhat less). For disabled students the values are significantly higher, $r = .83$ (Fall) and $r = .81$ (Spring) leading one to deduce using DIBELS ORF to predict ISTEP results is more useful for students with IEP's than those who are not considered disabled.

The same appears to be true for students whose native language is not English (ELL). Regression correlations between ORF and English/Language Arts ISTEP are $r = .85$ (Fall) and $r = .82$ (Spring). For these third grade students around 70% of their ISTEP performance can be predicted from their 2nd grade DIBELS ORF scores according the sample analyzed in this study. However, regarding the results for both disabled and ESL students, the N used for regression was 43 and 49 respectively, short of the minimum of 105 cases as recommended by Green (1991). Therefore, assumptions about how well these results can be generalized should be mitigated with consideration of the small sample.

An evaluation of regression results between ORF and ISTEP Math yields less noteworthy differences. Generally the correlations are around .60 for the Fall Math scores, and between .50 and .55 for Spring Math. The exceptions include the results from the subgroups of race, SES, and ELL. With regard to race, similar to findings pertaining to English/Language Arts, the strongest relationship between the scores belonged to non-white students, which were closer to the averages for the whole group. The ORF-Math linear regression statistics for white students were significantly lower than the others, $r = .53$ (Fall), and $r = .42$ (Spring). An equivalent difference was found for the SES category - students not considered as impoverished showed lower correlations.

The strongest relationship between ORF and ISTEP Math was found to be among students for whom English is a second language (ELL), and whose results declined the least from the Fall administration to the Spring. For this group Fall Math $r = .70$ (the highest among all the subgroups), and for Spring $r = .64$. From this analysis, it appears DIBELS ORF is a better predictor of ISTEP Math performance for students who are

generally considered academically challenged (non-white, poor, and have English language deficiencies) than students who are not considered at-risk. Again, consideration needs to be given to the small sample size for students in this group (non-white and ELL).

A general observation across the matrix shows the intercepts (or constants) tend to be lower values for non-white, impoverished, disabled, and ESL students. This is the point at which it is assumed students would score on ISTEP without regard to any postulated effect oral reading fluency might have on overall performance. As such, for the most part, regression coefficients (Beta values) seem higher across-the-board for these groups as well. From this inspection it is inferred that students who may be lower academically show more distinct gains in achievement as their oral reading ability improves, at least to the extent of how it is measured on DIBELS ORF and ISTEP. Conversely, the regression scores of normally higher achieving students who are more likely to belong to the subgroups female, white, and not disabled or not poor, show weaker correlations between reading fluency and performance on ISTEP. Overall, though, from this study it has been established there is a relationship between how well a student does on the DIBELS Oral Reading Fluency test (end of second grade) and later performance on ISTEP English/Language Arts and Math (third grade). There are, however, differences in the strength of that relationship among different types of students and between the two tests (English/Language Arts and Math), and the two test administrations (Fall and Spring).

Model Evaluated

As reported, the sample of data provided for this report included DIBELS and ISTEP scores from 306 students, each of which belonged to one or more the following subgroups – male or female, from poverty or not, disabled or not, white or nonwhite, and being an English Language Learner or not. The basic statistical model chosen for the study was simple regression, which was applied to the full set of data as well as each subgroup individually. However, to affirm whether or not this was the best way to evaluate the results and the possible effects the subgroups' ORF scores may have regarding the predictability of ISTEP, a hierarchical model was applied. The only set of scores examined with this model was DIBELS ORF and the Spring English/Language Arts (See Table 12).

Table 12

Hierarchical Regression Results for Spring ISTEP Ela

Model	<i>R</i>	<i>R</i> ²	Adj. <i>R</i> ²	Std. Error of Measurement	Change Statistics		
					<i>R</i> ² Change	<i>F</i> Change	<i>df</i>
ORF	.72	.52	.52	47.17	.52	332.99	1
ORF Race	.74	.55	.55	45.93	.03	17.58	1
ORF Race SES	.75	.56	.55	45.67	.01	4.45	2

Note. $p < .05$ for all models; gender, disability, and ELL were *ns*.

The *F* statistics of three of the subgroups, gender, disability, and ELL were found not to be significant at $p < .05$, and were therefore removed from the equation. It is unlikely the scores from these students affected the overall results of the regression analysis. However, SES and race seem to have a very minor effect. When only the independent variable ORF is considered, $r = .72$; with ORF and race $r = .74$, and with

ORF, race, and SES $r = .75$. In this model only an additional 3 - 4% of the variance between ORF scores and ISTEP can be explained by the effects of race and socioeconomic status. Evaluating only these statistics, one might conclude little or no differences between the subgroups with regard to the relationship there is between DIBELS and ISTEP. Three variables were excluded due to having no statistical significance.

Accepting the null hypothesis that no significant differences are present with regard to the unique characteristics that exist among subgroups in a sample is a Type VI error, and possibly due to statistical models that do not match the research design (Tracz, et. al., 2005). The results of hierarchical model presented suggest there is likely to be very little, if any at all, significant difference in the correlation between ORF scores and ISTEP results for the subgroups in the context of the full sample. Each demographic was examined simultaneously, therefore the effects of gender, SES, and race were incorporated in the overall results. For instance, a student may be both female and non-white; which characteristic most affected the outcome? Even by controlling for each demographic with a stepwise approach to hierarchical regression, the results are what Tracz called a “adjusted scores” and may not fully mean what the researcher thinks it does (Tracz, et. al., 2005). Whether hierarchical or simple regression techniques are used may be arguable among researchers; which process yields the most valid findings? For this study the researcher chose a simpler approach with simple linear regression, or correlation, with the understanding the relationships found in the different subgroups may look different in a covariance model. From this study and the technique used results

imply ORF is a better predictor for students from poverty, students of color, students with disabilities, and for English Language learners (see Table 13).

Pass/Fail

Logistic regression models were applied to determine the predictability of DIBELS ORF for ISTEP with the ISTEP scale scores transformed into dichotomous variables 1 or 0, pass or fail. The cut score for English/Language arts was 404 and for Math it was 393. Therefore, scale scores 404 and above were transformed into 1 for E/LA to signify pass, and less than 404 to 0 to signify fail. Likewise, for Math, scale scores 393 and above were coded 1, and scale scores 392 and below were coded 0. An analysis was applied to each of the four dependent variables, Fall and Spring ISTEP English/Language Arts and Math. The first statistics shown are from the base model of the logistic regression produced from SPSS. For the first classification model, the default was to pass.

Table 13

Dependent Variable Classifications

Test	Fail	Pass	Pct Correct
Fall 2008			
Ela	188	118	61.4
Math	176	130	57.5
Spring 2009			
Ela	218	88	71.2
Math	222	84	72.5

Table 14

Variables in the Equation

Test	<i>B</i>	S.E.	Wald	<i>df</i>	Sig.	Exp(B)
Fall 2008						
Ela	.47	.12	15.73	1	.000	1.59
Math	.30	.12	6.86	1	.009	1.35

Spring 2009						
Ela	.47	.12	15.73	1	.000	1.59
Math	.30	.12	6.87	1	.009	1.35

Without consideration of the predictor variable, DIBELS ORF scores, the initial model predicts 61.4% and 72.5% passing rate on ISTEP. For each of the dependent variables, the statistics are significant with $p < .01$. The only variable not included in the initial model is the predictor variable, which is also found to be statistically significant (using chi-square), meaning the addition of this variable to logistic regression will significantly affect the predictive power of the model.

r Square values for the Cox & Snell and Nagelkerke are analogous to r^2 findings in linear regression (Miles, 2004), therefore a reasonable fit of the data to the logistic regression model is shown albeit with lower values for Math from both tests. Typically, the Nagelkerke R Square calculations are primarily considered (Miles, 2004).

Table 15

Model Summary

Test	-2 Log likelihood	Cox & Snell r Square	Nagelkerke Square
Fall 08 Ela	280.74	.34	.46
Fall 08 Math	273.63	.25	.36
Spring 09 Ela	207.00	.41	.58
Spring 09 Math	334.916	.236	.317

There is a considerable improvement in the predictability of ORF on ISTEP from Fall to Spring (Table 15).

In the initial model (not shown) the -2 log likelihood statistic was 334.916 when only the constant was considered. Values for -2 log likelihood for all the ISTEP values

are less after introducing the dependent variables, indicating the second model is predicting the outcomes more accurately. Also, after all the variables were entered into the equation the Omnibus Test of Model Coefficients using Chi-square, which is similar to the t-test in linear regression, found the model to be significant at the $p < .01$ level, meaning that it is unlikely chance alone would account for the predictions better than the model did.

The next phase of the analysis is to add the ORF scores to the calculations to test whether or not there is, in fact, an improvement in the prediction. The following tables summarize the results.

Based on the introduction of the outcome variables prediction increased for passing all ISTEP tests. The probabilities for DIBELS to predict passing English/Language arts increased from 61.4 to 78.4 (Fall) and from 71.2 to 85.0 (Spring). For Math the predictions grew from 57.5 to 72.9 (Fall) and from 72.5 to 80.7 (Spring). As with correlations examined earlier, DIBELS ORF is a better predictor for passing

Table 16

Classification Table after independent variable (ORF) is applied

<i>Test</i>	<i>Predicted</i>		<i>Pct. Correct</i>
	<i>Fail</i>	<i>Pass</i>	
Fall 08 Ela			
Fail	78	40	66
Pass	26	162	86
Overall Pct.			78
Fall 08 Math	80	50	
Fail	80	50	62
Pass	33	143	81
Overall Pct.			73
Spring 09 Ela			
Fail	57	31	65
Pass	15	203	93
Overall Pct.			85
Spring 09 Math			
Fail	39	45	46.4
Pass	14	208	93.7
Overall Pct.			80.7

English/Language Arts than for Math. A marked difference in the logistic regression model in comparison to linear regression is the percentages predicted for passing are notably higher for the Spring ISTEP tests. In the linear model, r and r^2 scores were similar, especially for Language Arts (See Table 16).

The “Variables in the Equation Table” displays the estimated coefficients for the predictor variables, which are the B values and are similar to B values in linear regression. In this model the interpretation is one change in the unit of the predictor score will increase the logit of the dependent variable by the amount of the coefficient.

Table 17

Variables in the Equation

	<i>B</i>	<i>S.E</i>	<i>Wald</i>	<i>df</i>	<i>Sig.</i>	<i>Exp(B)</i>	<i>95% C. I. Exp (B)</i>	
							<i>Lower</i>	<i>Upper</i>
Fall 2008								
ORF	.049	.01	70.57	1	.000	1.05		
Ela Constant	3.63	.50	52.73	1	.000	.027	1.04	1.06
ORF	.034	.01	56.80	1	.000	1.04	1.03	1.04
Math Constant	2.63	.41	41.86	1	.000	.072		
Spring 2009								
ORF	.066	.01	71.86	1	.000	1.07	1.05	1.08
Ela Constant	4.18	.59	49.71	1	.000	.015		
ORF	.039	.01	58.62	1	.000	1.04	1.03	1.05
Math Constant	2.09	.40	27.92	1	.000	.12		

This can be illustrated by the following equations (Miles, 2004):

$$\text{Logit}(\text{pass}) = B \times \text{score} + \text{constant}$$

$$\text{Odds ratio} = \exp(\text{logit}(\text{pass}))$$

$$P = \text{odds} / (1 + \text{odds})$$

By applying the equation to the data produced by the logistic regression model, the following probabilities occur (Miles & Shevlin, 2004).

Table 19

Probabilities Passing ISTEP from ORF scores

<i>Test</i>	<i>p for 80 cwpm*</i>	<i>p for 90 cwpm*</i>	<i>p for 100 cwpm*</i>
Fall Language Art	.57	.69	.78
Fall Math	.52	.61	.68
Spring Language Arts	.75	.85	.92
Spring Math	.74	.81	.86

*DIBLES ORF – Number of words read correctly per minute

An oral reading fluency score on DIBELS of 90 correct words per minute or more is considered “low risk,” or reading at the expected level. Therefore, an ORF score of 90 is used in the calculations to yield probabilities for passing ISTEP (middle column). Probability values for reading ten less words per minute (80) as well as 10 more words per minute (100) are also shown. Again, according these results the likelihood of students passing ISTEP is less for the Fall test in both Language Arts and Math. From this analysis it appears DIBELS ORF is a better predictor for ISTEP tests in the Spring, almost a year after ORF is taken. Also, for students who achieve a reading score of 100, this model suggests a very strong probability they will pass both parts of ISTEP in the Spring.

Summary

Linear, multiple, and logistic regression models were used to evaluate the relationship between DIBELS ORF scores and ISTEP performance. Through all the examinations a relationship was found and it is determined oral reading fluency as measured by the DIBELS test can be a predictor to how well students will do on ensuing ISTEP tests. The relationship was clearly stronger for English/Language Arts than Math, but significant correlations were also found between reading and math scores. When the scores were segregated into categories of gender, race, SES, disability, and ELL some differences in the correlations did emerge. ORF seems to be a better predictor for ISTEP results for students who are generally considered at-risk than those students who are not. Logistic regression tested the possibility of using DIBELS ORF to predict whether or not students would pass ISTEP; the finding being that it can. Aside from a small concern

about having a normally distributed sample for the study, generally all statistical tests performed showed significance at the $p < .01$ level and the null hypothesis was rejected. Therefore, it is likely the results found in this study are also represented in the population.

CHAPTER 5

Conclusion

Introduction

The primary goal of this study was to determine if there is a relationship between students' oral reading fluency score on DIBELS ORF at the end of the second grade and those same students' scores on ISTEP the following year. ISTEP is the Indiana statewide test which meets the requirement of federal law mandating the annual assessment of student academic achievement. In Indiana this test is administered to all students in grades three through eight, the results of which are used to determine the strength of a school's academic program and provide a designation of whether or not that school is meeting "Adequate Yearly Progress (AYP)." The stakes are high, for if schools fail to meet progressive standards they can receive negative consequences in the form of sanctions from the State Department of Education as well as the federal government. Loss of local control with regard to decision-making about educational programming as well as a loss or redistribution of revenue can be retributions for failing to meet required academic standards. Since ISTEP is a comprehensive, large annual summative exam and is generally given only once per year, the ability to have a means to predict students' future performances provide educators the ability to take corrective action and implement

instructional interventions to insure better results on the statewide test and avoid reprisals.

State education policy in Indiana places a strong focus on the ability of students to read fluently and comprehend at the end of the third grade. Those students failing to show mastery on a state imposed reading test are required to take expensive remedial summer reading programs and/or be retained in the second grade. Therefore, the proportion of students failing to meet reading standards at this level will be attributed to the failure of a school to perform adequately leading to more possible sanctions including the label of a "failing school." Certainly the fundamental goal for all schools is to properly educate its students and not be overly concerned about government imposed tests, but state and federal accountability laws force the issue into the minds of the public and the stigma associated with negative reports in the media cannot be avoided. Such an impetus understandably affects the emotional and cultural status of a school's faculty and administration generally in a negative way. For that reason by choosing to find tools to help predict students' performances on ISTEP should bring the focus back to meaningful instruction and away from test and AYP anxiety. With the proliferation of the use of DIBELS in schools across Indiana as a part of a statewide program called "Wireless Generation," it seemed appropriate to examine the relationship between DIBELS ORF and ISTEP for this study. If there is, in fact, a strong correlation between reading fluency and ISTEP scores then educators will have viable means to guide instruction and take appropriate intercessory actions on a timely basis.

Overall Findings

Sample data was collected from four different elementary schools in different parts of Indiana and included scores from students in a large corporation, a medium sized corporation, and one small corporation. The total number of cases of 306 were comprised of students' end-of-second grade DIBELS ORF scores and these same students' scores from ISTEP the following year in the third grade. The test results were drawn from archived data from the years 2008 and 2009 since that was the only year ISTEP was administered twice as part of a plan to transition from a Fall test to a Spring test. The assumption was having the advantage to compare ORF scores to two sets of ISTEP results would provide additional depth to the final analysis of the statistical findings. ISTEP is composed of two parts, English/Language Arts and Math; the DIBELS oral reading scores were compared to the results of both parts. Not only was an attempt made to examine the possible correlation between ORF and Reading, but to see if a relationship existed for Math as well.

The Relationship

Since the objective was prediction, linear regression was used as the first statistical model. The total group of 306 students was analyzed as well as each of the subgroups segregated by gender, race, SES, disability, and ELL. Overall the regression statistics showed a significant relationship does exist between second grade DIBELS ORF and third grade ISTEP and that there is utility pertaining to prediction. This is true for Math as well as English/Language Arts, although not as strong. The r value for English/Language Arts was .76 for the Fall administration of ISTEP and for Spring it was .72. The fact that this correlation held fairly constant across both the Fall and the Spring

tests, separated by five months (September - March), infers the sense that DIBELS ORF is a valid predictor of ISTEP Reading/Language Arts performance. In short, according to r Square values for both test administrations ($r^2 = .57$, Fall; $r^2 = .52$, Spring) over 50% of the variance in ISTEP E/LA scores can be explained by how well students performed on DIBELS ORF at the end of the second grade. Certainly other dynamics account for overall performance on achievement tests, but if one variable such as reading fluency ability can be implied to have an effect equal to or more than half, it is an important consideration.

From the regression statistics it was found the correlation between DIBELS ORF and ISTEP Math was not as pronounced as with English/Language Arts, however a relationship did exist. Also, the strength of the relationship appeared to decline from the Fall administration of ISTEP to the Spring. The regression correlation between ORF and Fall Math was .61, and for Spring it was .54. r Square values equal .37 (Fall) and .29 (Spring). Some of the variability of the regression results might be explained because the ISTEP tests were not identical; some changes were made to reflect the expected achievement level for third grade students who had undergone almost a full year of additional instruction. However, the same is true for the English/Language Arts sections of ISTEP as well, and that correlation seemed to hold fairly constant. Also, it is intuitive that cognition necessary for literacy related questions and problems can be different for numeracy. DIBELS is a reading test and the assumption can be made to the extent oral reading fluency is related to Math achievement is due to the need to be able to read the problems on the test.

The Subgroups

Separate regressions were run for each of the subgroups to determine if segregating the results would yield additional information about the usefulness of using DIBELS ORF to predict ISTEP scores. Groups analyzed included gender (male or female), race (white or non-white), students eligible for lunch and textbook assistance (yes or no), disability (with or without), and students whose native language is not English (yes or no). Regression calculations were applied to each category separately from the others. The intent was to isolate a single dependent variable with a certain characteristic in order to make comparisons from one group to the other, as well as a comparison to the whole group results.

The general findings show the scores from those student groups considered the least at-risk for academic failure, white, female, not poor, not disabled, and have English as their native language had lower correlations between ORF and ISTEP than the scores from students of the most at-risk groups. Though the relationship between the performances on both tests is significant in most cases for all groups, for students belonging to categories suspected to be academically challenged DIBELS ORF may be a better predictor of performance on ISTEP. For ISTEP English/Language Arts, the strongest regression scores were for students with disabilities ($r = .83$ Fall, $r = .82$ Spring) and English language learners ($r = .85$ Fall, $r = .81$ Spring). Though the small number of these cases for regression analysis needs to be considered, the correlations are rather strong.

A comparison of math correlations show marked differences for students from poverty and students whose native language is not English. The relationship between

DIBELS ORF and ISTEP scores from impoverished students is notably higher ($r = .62$ Fall, $r = .53$ Spring) than scores from students who are not considered poor ($r = .49$ Fall, $r = .41$ Spring). A similar variation in correlations of scores was found between ELL students ($r = .70$ Fall, $r = .64$ Spring) and English speaking students ($r = .58$ Fall, $r = .50$ Spring). Again, the small sample sizes need to be taken into account. With these results, one might explain about 28% - 49% of ISTEP MATH scores for students of poverty and non-proficient in English can be attributed to oral reading fluency performance on DIBELS. For students who are not considered poor and non-ELL, as well as the whole group, the percentages range from 17% - 38%. In summary, the predictability of DIBELS ORF on Math ISTEP for third graders is of moderate or minimal strength, depending on the subgroup considered

Research Questions Answered

In addition to analyzing the overall relationship between DIBELS ORF and ISTEP and the predictive power of oral reading fluency scores, the following research questions were posed and answered.

- Do third grade students who read at or above the “low risk” designation of DIBELS ORF at the end of their second grade year pass the Language Arts portion of ISTEP in third grade?

With regard to the sample obtained for this study, on the Fall test a total of 188 students passed (out of 306,) indicating an overall passing rate of 61.4%. Of those students passing, 54 students scored below the threshold of 90 correct words per minute

on DIBELS. There were 19 scores above 90 correct words per minute on ORF which did not qualify for a passing score on ISTEP. The results of the Spring ISTEP test show 218 students passing, a rate of 71.2%. In this case 75 students passed who had scores of 89 or less on DIBELS, and 15 students failed who scored 90 or more on ORF. When adjusting for the scores of students failing DIBELS but passing ISTEP, the remainder is 135 students passing Fall ISTEP English/Language Arts and 155 passing in the Spring. These numbers account for students who met the minimum acceptable score on DIBELS and also passed ISTEP, rates of 44.1% and 50.6% for the Fall and Spring tests respectively.

According to the results of logistic regression analyses, for Fall ISTEP the statistical model predicted 208 students to pass, 288 did; for the Spring test, 208 were predicted to pass, in actuality 218 did. The model purports a successful prediction for passing ISTEP Language Arts at 86.2% for the Fall test and 93.1% for the Spring. In consideration of both the descriptive and inferential statistics, it does appear students who score at the “low risk” designation on DIBELS ORF at 90 words correct per minute at the end of their second grade year are more likely to pass the English/Language Arts portion of ISTEP as third graders.

- Do third grade students who read at or above “low risk” designation of DIBELS ORF at the end of their second grade year pass the Math portion of ISTEP in third grade?

In Math 271 of 306 students passed ISTEP in the Fall (88.6%), and 291 passed in the Spring (95.0%). Adjusting for students who did not meet the DIBELS ORF standard, but

did pass ISTEP, the numbers are 132 passing Fall Math and 147 passed in the Spring.

Regarding students who met or exceeded the score of 90 on DIBELS, 35 did not pass in the Fall, and 15 failed in the Spring. When the data set was applied to logistic regression, the model predicted passing rates of 81.3% for Fall Math and 93.7% in the Spring.

Similar to English/Language Arts, it does appear one can use DIBELS ORF test results to predict whether or not a student will pass ISTEP Math.

- Can the number of correct words per minute as scored by DIBELS ORF at the end of the second grade year be used as a predictor of whether or not that student will pass the spring ISTEP (both Language Arts and Math) in the third grade?

By using Beta and constant values from logistic regression analyses, as well as selected oral reading fluency scores, probabilities were calculated for passing ISTEP. If a student reads at a rate of 90 words correct per minute, the probability he or she will pass ISTEP Language Arts in the Fall is .69 and .81 in the Spring. If the DIBELS score is increased by 10 points (to 100), the probabilities grow to .78 (Fall) and .92 (Spring). Conversely if the reading score is decreased by 10 points (80), probabilities decrease as well to .57 in the Fall and .75 in the Spring. Though the threshold of 90 correct words per minute on DIBELS ORF is considered “low risk” for reading, and thus for academic achievement, oral reading scores of 100 or more provide a solid basis from which to predict ISTEP English/Language Arts performance.

The same calculations were applied to ISTEP Math scores as with English/Language Arts. Results yielded similar findings. At the standard of 90 correct words per minute, the probability for students passing Fall Math was .61 and the probability for passing Spring Math was .81. When the reading score increased to 100, the probabilities increased as well – to .68 (Fall) and .86 (Spring). The negative effect was larger with scores on the ORF ten points less (80); probabilities fell to .52 (Fall) and .74 (Spring).

In short, the higher the DIBELS score, the more likely a student is to Pass ISTEP in both English/Language Arts and Math.

- Is the relationship, if any, between the end of the second grade DIBELS ORF and third grade ISTEP scores in the Fall similar to a relationship between these scores in the Spring?

Linear and logistic regression analyses both indicate a relationship between scores on second grade end-of-year DIBELS ORF and third grade ISTEP scores, however the image of that relationship varies somewhat between the Fall and Spring ISTEP administrations and between the types of analyses. Correlations found from linear regression (r) are shown to be similar between the Fall and Spring, values of .76 and .72 respectively. For Math the relationship varies from $r = .61$ for the Fall and $r = .54$ in the Spring, a notable decline in strength. With logistic regression, in an attempt to assess the predictive power of DIBELS ORF on passing ISTEP, a different picture occurs. The results of these statistical tests show the predictability of using oral reading fluency

scores for ISTEP is stronger for the Spring administrations in both subjects, Language Arts and Math. The logistic regression model indicates 90%+ success rate of predicting passing Spring ISTEP from DIBELS ORF scores, an improvement over the predictions for the Fall tests by approximately 7% (ELA) and 32% (math).

The two measures, linear and logistic regression models, evaluate the data differently and answer different questions. Linear regression tests to see if the ORF and ISTEP scores are related, and if that relationship, when values are plotted on a graph, shows a line and slope that increases as both scores increase. Beta values, or regression coefficients, from these tests show that it does, though with some dissimilarity between the two ISTEP subjects and seasonal administrations. Enough information yielded by linear regression is provided to discern a degree of prediction, that a student's reading score on DIBELS can be used to predict a subsequent score on ISTEP. Logistic regression, on the other hand, seeks to predict the answer to a simple question as to whether or not, based on an ORF score, students will either pass ISTEP or not.

- Is the relationship between DIBELS ORF scores and ISTEP performance similar or different between students from differing demographic subgroups such as gender, race (white, nonwhite), students with disabilities and those without, students in poverty and those who are not, and English Language Learners and those who are not?

Generally, the relationship between DIBELS ORF and ISTEP are stronger for students who are normally considered learning challenged, such as students with

disabilities and English Language learners. Correlations between end-of-second grade oral reading scores and Spring ISTEP English/Language Arts scores are over .80 for disabled and ELL students. Correlations between DIBELS and ISTEP Math scores were also higher for these students as well. The lowest correlations between any of the scores were for female students, white students, and students not economically disadvantaged, though the relationship between DIBELS and ISTEP was still moderate to fairly strong (r ranged .60 to .69 for Spring ELA, and .41 to .55 for Spring Math).

General Summary of the Conclusion

The findings of this study compare to the results of other studies reviewed in an earlier section of this report. Similar to the state tests for Florida (FCAT), Oregon (OSA), Ohio (OPT), Pennsylvania (PSST), The Iowa Test of Basic Skills, and the North Carolina End of Grade Test, the relationship between DIBELS and the Indiana test (ISTEP) appear to have correlations of around .70 or more. This study also infers a relationship between oral reading skills and math performance as well with somewhat less strength with r values between .54 and .61; scores are a little lower than Aiken reported in his study regarding oral reading and math performance on the California Achievement Test (CAT) (Aiken, 1992). Adding to the body of work on this topic this study furthers validates the theorized connection between reading ability and achievement as it is measured by standardized tests.

Conclusions drawn about using DIBELS ORF as a predictor for ISTEP as a result of this study include the reasonable ability to forecast a student's scale score on both portions of ISTEP, and using a DIBELS score to attain a sense of probability a student is likely to pass or not. Vander Meer reported students' scores of 110 or more on oral

reading fluency resulted in a 72% passing rate on the Ohio Proficiency Test (Vander Meer, et. al., 2005). Results from logistic regression analyses on the data sample for this study found the probability for passing ISTEP English/Language Arts to be between .78 and .92 for students achieving a score on ORF at 100 correct words per minute. With this ORF score, for Math the probability values were between .68 and .86, from a somewhat lower reading score than reported by Vander Meer. For the Oregon test, it was reported an oral reading score of 110 words per minute established the threshold where students were likely to meet or exceed expectations on the OSA (Good III, et. al., 2001). Nonetheless, though quantitative differences are apparent on exactly what DIBELS ORF score predicts the best outcome, the notion that oral reading fluency ability may directly affect performance on statewide tests is apparent, and is a significant notion for educators to hold as they prepare students to take these tests.

Judging from the results of linear regression analysis on the subgroups in this study, DIBELS ORF appears to be especially useful for predicting ISTEP performance for students normally considered learning challenged. Contrary to Roehrig's findings, that hierarchical regression analyses did not find significant interaction across the demographic subgroups, or were significant contributors to prediction, this study did find noteworthy differences in simple linear regression correlations. Scores from students of color, disabled, or English Language Learners showed higher correlations than did their counterparts. However, the sample size was small and was extracted from the larger group. A single linear regression model was applied to each subgroup, unlike the hierarchical model used by Roehrig (2007). Also, normality tests of the subgroups analyzed in this report did not fall within acceptable standards and this fact should be

considered. However, when a fairly normally distributed sample as was used in this study is broken down and specific pieces of it are examined in isolation, it's understandable the parts may not mirror the overall population. As such, this researcher will give some credence to the regression analyses of the subgroups and the differences between them.

The most basic question this study sought to answer was whether not students' oral reading fluency scores at the end of the second grade could be useful in predicting ISTEP scores for these same students in the third grade, a year later? The simple answer seems to be "yes." There clearly are no "magic bullets" in education; the behavior of learning is sophisticated and complex. A host of variables contribute to a student's development in life, but it does seem that the ability to read well is a fundamental skill for success in school.

Discussion

To assess adequate yearly progress for schools, Indiana initiated a new rating scale in 2011. Based on ISTEP results ratings for schools changed from the five designations of exemplary progress, commendable progress, academic progress, academic watch, and academic probation to letter grades A, B, C, D, and F respectively. Legislators, as well as the State Superintendent of Public Instruction, often asserted in media reports the need for a labeling system the public could understand and thereby be better informed about the status of their local schools. Certainly a school receiving a letter grade of F is more easily understood than the designation of "academic probation." Though clarity was achieved about how well a school was doing or not doing, the anxiety felt by educators and students was also heightened. The emotionally charged nature of

letter grades, especially bad ones, can understandably increase the worry about ISTEP performance; such anxiety may very well spur increased efforts to “teach to the test,” and other attempts to prepare students for ISTEP by using low level, rote memorization, and content focus teaching methods. The high stakes nature of such tests negatively impacts the quality of instruction and prompts excessive amount of time for test preparation (Shepard, et. al, 1991).

A study by Monica Osburn at the University of Arkansas examined the stress and anxiety that accompanies standardized test-taking, as well as attitudes held for such by parents, teachers, and students. She found parents were concerned about the pressure teachers are under with regard to testing, and that for students who performed poorly on such tests, the anxiety levels were higher (Osburn,etal, 2004). Osburn referred to other studiesby Bernaur and Cress (1997) and Paris (1992) that shed doubt on the value of standardized test results as a measure of student learning, and one by Murry (1998) that questioned the undo emphasis on these test results (as cited by Osburn, et. al, 2004). Additional research implies graduation rates suffer; students, especially those who are economically disadvantaged or have low GPA’s, are more likely to drop out of school when passing a standardized test is a requirement for graduation (Reardon, 2002). Such tests also decrease student motivation and increase retention (Amrein, 2003).

Yet, those in state and federal government responsible for the development of education policy seem to ignore the literature and research as it pertains to the effects such a strong emphasis on standardized testing has on student learning and school culture. In Indiana, by affixing a letter grade to schools with respect to test results seems only to have exacerbated the problem. Hence, the natural tendency for educators to

respond to community, media, and political pressure to avoid a low grade is to deviate from what is generally accepted as best teaching practices. Such methods include a focus on direct instruction, the development of higher level thinking skills about the subject matter which students learn, and the relevance and usefulness of the skills they acquire.

In this light, the use of a formative assessment such as DIBELS ORF can aide teachers in their attempt to foster strong literacy skills within their students; as a result these same students will ultimately perform better on a high stakes test such as ISTEP. The anxiety about adequate yearly progress and the likelihood of receiving a poor grade for the school can be mitigated by more emphasis on teaching and improving students' abilities to read fluently, and less emphasis on "test prepping." The implication regarding the results of this study that examined the relationship between oral reading fluency and ISTEP, both for the English/Language Arts and Math sections, is considerable for second and third grade students. The first experience students in Indiana have with regard to a high stakes test is in the third grade, so it seems sensible for teachers at this level and lower grades to be more cognizant of and attentive to the development of strong reading skills within their students throughout the school year, and have less concern about ISTEP in the Spring. This appears to be especially true for students from economically distressed households, special education students, students of color, and students who are English Language Learners. If oral reading fluency improves, so does the probability of a student passing ISTEP (See Figures 1 and 2).

Figure 1

Probabilities of passing ISTEP English/Language Arts with a certain ORF score

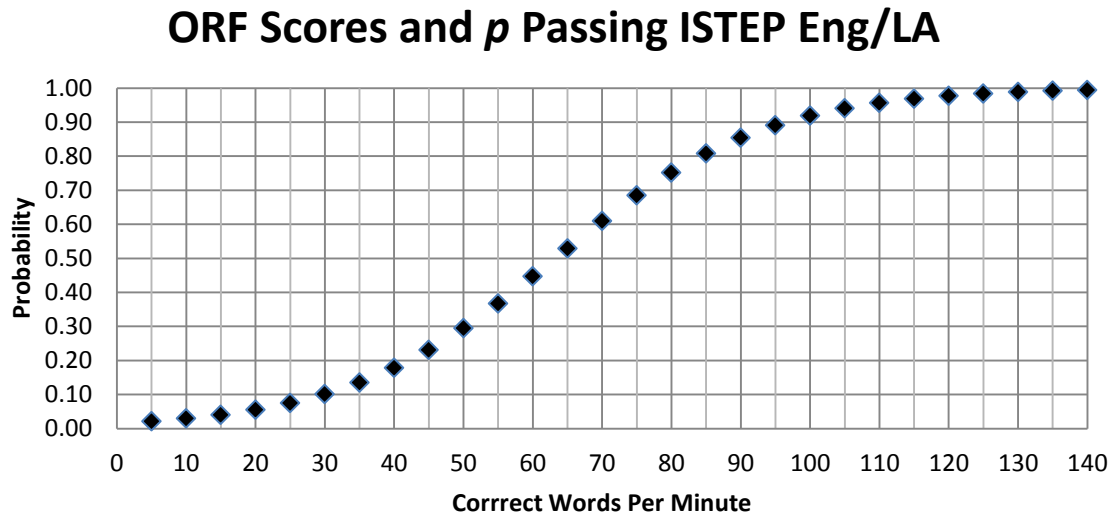
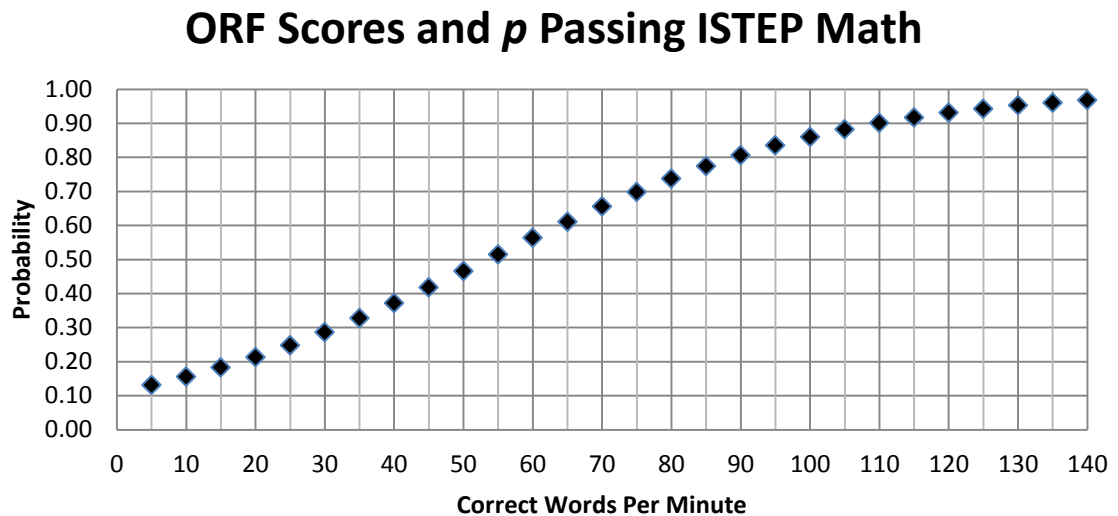


Figure 2

Probabilities of passing ISTEP Math with a certain ORF score



The graphs were constructed from results of logistic regression models that evaluated the usefulness of using DIBELS ORF scores to predict ISTEP performance. As reported earlier, the relationship between the two assessments is significant, therefore

by improving oral reading fluency, there is a good chance ISTEP results will improve as well in both Reading/Language Arts and Math. The “low risk” designation for DIBELS ORF at the end of the second grade is 90 or more correct words per minute, which, according to the graph, infers a probability of about .85 of passing ELA the following year, and about a .80 probability for passing Math, which are pretty good odds. However, by working with students throughout the year to develop better oral reading fluency skills, if the ORF scores increase by 10 correct words per minute, those probabilities increase to .90 and .88. Once a student is reading fluently at rates of 120 words or more, probabilities begin to approach the maximum likelihood of passing both tests.

For all students, especially those who are learning challenged, direct instruction (teacher directed with references to text and emphasis on students self-correcting), sequencing of tasks, repetition, segmentation (breaking down skills into smaller parts), directed questioning and responses, the use of technology, and teacher modeling of problem solving are effective teaching methods (Swanson, 1999). Also, with regard to fluency, lessons in phonemic awareness and phonics are also important to early readers.

By periodically using DIBELS ORF to assess a student’s reading fluency, when there are problems interventions can be provided on a timely basis; such remedial lessons can be drawn from the chronology of how children acquire the ability to decode and cite words. For instance, being able to distinguish between the sounds of language (phonemic awareness) and learning to apply the letter-sound relationship to speech (phonics) is hierarchical, the first is dependent on the latter. As students are able to discern different sounds it is easier for them to make sense of letters and the particular sounds attached to

each one. As phonics skills are enhanced students are better able to decode words and improve vocabulary – essential to growth in oral reading proficiency.

These are things teachers can choose to do, every day, every hour, to assist their students to become better readers. It can be strenuous work, obstacles and frustrations occur, but with careful guided instruction and practice students will get better, even those with cognitive impairments or other issues interfering with their learning. The choice to engage in meaningful, relevant instruction, guided by frequent formative assessments such as DIBELS ORF, should help dispel the urge to be overly concerned about ISTEP. As implied by this report, while students become better readers in the second grade, they also improve their chances significantly at being able to pass ISTEP in the third grade. Though the full weight of school accountability cannot be eliminated entirely, it can be lessened when teachers are aware of the tools they have and can use to inform their instruction and monitor their students' learning.

REFERENCES

- Amrein, A., & Berliner, D. (2003). The effects of high stakes testing on student motivation and learning. *Educational Leadership*, 60(5), 32-38.
- Barger, J. (2003). *Comparing DIBELS Oral Reading Fluency indicator and the North Carolina end of grade reading assessment* (Technical Report). Asheville, NC: North Carolina Teacher Academy.
- Beaver, M. (1998). *The effects of a language arts block on student achievement*. Report No. CS 216 351). Vigo County School Corporation, Terre Haute, IN.(ERIC Document Reproduction Service No. ED 420 062)
- Burk, A. (1998). *A literature based basal reader's effect on student achievement: an evaluation of literature based textbook adoption at Terre Town Elementary*. (Report No. CS 216 349). Vigo County School Corporation, Terre Haute, IN.(ERIC Document Reproduction Service No. ED 420 060)
- Chard, D. J., Vaughn, S., & Tyler, B. (2002).A Synthesis of Research on Effective Interventions for Building Reading Fluency with Elementary Students with Learning Disabilities. *Journal of Learning Disabilities*, 35(5), 386-406.
- Cole, H., Hulley, K., & Quarles, P. (2009). Does assessment have to drive the curriculum? *Forum on public policy online*. 2009(1).
www.forumonpublicpolicy.com
- Dynamic Measurement Group (2004). *DIBELS 6th Edition Technical Adequacy Information*. <http://dibels.org/pubs.html>.
- Field, A. (2009). *Discovering statistics using SPSS*. Sage publications. London.

- Goffreda, C. T., & DiPerna, J. C. (2010). An empirical review of psychometric evidence for the Dynamic Indicators of Basic Early Literacy Skills. *School Psychology Review, 39*, 463-483.
- Good III, R. H., Simmons, D. C., & Kame'enui, D. J. (2001). The importance of decision making utility of a continuum of fluency based indicators of foundation reading skills for third grade high stakes outcomes. *Scientific Studies of Reading, 5*(3), 257-288.
- Green, S. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavior Research, 26*(3), 499-510.
- Indiana Department of Education. Retrieval date from (2008-2009) www.doe.in.gov
- Jewell, J. (2008). Influence of reading attitude on reading achievement: a test of the temporal interaction model. *Psychology in the Schools, 45*(10).
- Jimerson, S. R. (2001). Meta-Analysis of Grade Retention Research: Implications for Practice in the 21st Century. *School Psychology Review, 30*(3), 420-37.
- Jordan, N.C. (2007). Do words count? Connections between mathematics and reading difficulties. In Berch, D. B. and Mazzocco, M. M. (Eds.). *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (Chapter 6). Baltimore, MD: Paul H. Brookes.
- Lesnick, J., George, R. M., Smithgall, C., & Gwynne, J. (2010). *Reading on grade level in third grade: how is it related to high school performance and college enrollment (Technical Report)?* Chicago, Illinois. Chapin Hall at the University of Chicago.

- Martinez, R., Aricak, O., & Jewell, J. (2008). Influence of reading attitude on reading achievement: A test of the temporal-interaction model. *Psychology in the Schools*, 45(10), 1010-1022.
- Miles, J., & Shevlin, M. (2004). *Applying regression and correlation: a guide for students and researchers*. Sage Publications. London.
- Osburn, M., Stegman, C., Suitt, L., & Ritter, G. (2004). Parents' perceptions of standardized testing: Its relationship and effect on student achievement. *Journal of Educational Research & Policy Studies*, 4(1), 75-95.
- Paleologos, T. M. (2005). *The effectiveness of DIBELS oral reading fluency as a predictor of reading comprehension for high and low income students (Dissertation)*. Auburn University.
- Public Law 109. Reading Deficiency Remediation Act, Ind. Code 20-32-8.5
- Riedel, B. W. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first grade students. *Reading Research Quarterly*, 42(4), 546-562.
- Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2007). Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology*, 46(3), 343-366.
- Rutherford-Becker, K. J., & Vanderwood, M. L. (2009). Evaluation of the relationship between literacy and mathematical skills as assessed by curriculum based measures. *Contemporary School Psychology*, 14, 23-34.
- Salinger, T., Mueller, L., Song, M., Jin, Y., Zmach, C., Toplitz, M. & National Center for Education Evaluation and Regional Assistance (2010). Study of teacher

- preparation in early reading instruction. NCEE 2010-4036. *National Center For Education Evaluation And Regional Assistance*.
- Schilling, S. G., Carlisle, Joanne F., Scott, S. E., & Zeng, Ji. (2007). Are fluency measures accurate predictors of reading achievement? *Elementary School Journal*, 107(5), 429-448.
- Shapiro, E. S. (2008). Use of a measure of reading comprehension to enhance prediction on the high stakes assessment. *Learning and Individual Differences*, 18(3), 316-328.
- Shepard, L., & Dougherty, K. (1991). *Effects of high-stakes testing on instruction* (Technical Report). National Council for Measurement in Education. Chicago, IL.
- Stage, S. A., & Jacobsen, M. D. (2001). Predicting Student Success on a State-Mandated Performance-Based Assessment Using Oral Reading Fluency. *School Psychology Review*, 30(3), 407-419.
- Swanson, L. (1999). Intervention research for students with learning disabilities. *The Executive Parent*, 29(9), 62-63.
- Tracz, S., Nelson, L., Newman, I., & Beltran, A. (2005). The misuse of ANCOVA: The academic and political implications of type VI errors in studies of achievement and socioeconomic status. *Multiple Linear Regression Viewpoints*, 31(1), 16-21.
- Vander Meer, C. D., Lentz, F. E. & Stoller, S. (2005). *The relationship between oral reading fluency and the Ohio proficiency testing in reading* (Technical Report). Eugene, OR: University of Oregon.

Wright, R. J. (2009). Methods for improving test scores: The good, the bad, and the ugly. *Kappa Delta Pi Record*, 45(3), 116-121.

APPENDIX

LETTER INVITING PARTICIPATION

Date, 2011
Principal's Name
School System
Address

Dear (NAME),

The purpose of this letter is to request your assistance in a study that I am doing for my dissertation through the Department of Educational Leadership at Ball State University. The study seeks to find out whether or not DIBELS ORF scores from the end of the second grade can be used to predict performance on ISTEP in third grade. Several studies have been done in other states regarding the relationships between DIBELS and their statewide exams and correlations have been found. It is my opinion that if such a relationship exists between DIBELS ORF and ISTEP, then teachers will have a means to predict a student's score on ISTEP ahead of time and can then take necessary steps to provide meaningful instructional intervention if needed.

If you are interested, I will need your assistance in acquiring the following information: DIBELS ORF scores expressed as "correct words per minute" from students in second grade classes for the spring test of 2008. In addition I would need the ISTEP scale scores for both the Language Arts and Math portions of ISTEP from both the Fall 2008 administration and Spring 2009 administration.

It is very important to note, I do not seek the names of students, nor should the data be provided that would compromise student confidentiality. I simply need a list of students, identified by either a random number or letter that shows their 2008 Spring DIBELS score and their corresponding scale scores on the Fall 2008 ISTEP and Spring 2009 ISTEP (both Language Arts and Math). Also, I am seeking demographic information about each participant which can be codes as follows:

1 for male, 2 for female, 3 for poverty (receives assistance for school lunch and textbooks), 4 for African American, and 5 for any other ethnicity (nonwhite).

If you decide to assist me with this study by providing the information I have requested, by signing below you attest that you do so without the disclosure of any student identities. If you have any questions please don't hesitate to call or email me.

Thank you so much for your assistance. If there is anything I can do to ease the effort required in compiling this information, I would be happy to help in any way I can.

Sincerely,

Gary L. Storie, Ed.S.
Ball State University

NAME OF PRINCIPAL